



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

공학박사 학위논문

거리기반 후보군 선정과 심층신경망을
결합한 다목적 히스토리매칭 연구

**Multi-objective History Matching Integrated with
Deep Neural Network and Distance-based Candidate Selection**

2017년 8월

서울대학교 대학원

에너지시스템공학부

김 재 준

거리기반 후보군 선정과 심층신경망을
결합한 다목적 히스토리매칭 연구

지도교수 강 주 명

이 논문을 공학박사 학위논문으로 제출함

2017년 6월

서울대학교 대학원

에너지시스템공학부

김 재 준

김재준의 박사학위논문을 인준함

2017년 7월

위 원 장 _____ (인)

부 위 원 장 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

초 록

이 연구에서는 심층신경망과 거리기반 후보군 선정법을 결합하여 다목적 히스토리매칭을 수행하고 신뢰도 높게 저류층의 생산추이를 예측하는 기법을 개발하였다. 기존의 앙상블기반의 히스토리매칭 기법은 입력자료들이 정규분포를 따르지 않는 경우에는 최적해의 신뢰성이 떨어진다는 한계점이 있다. 또한, 다목적 최적화 알고리즘은 목적함수가 많아지면, 모든 목적함수를 만족시키는 최적해가 제대로 수렴하지 못하는 문제점이 있다.

이 연구에서 제안한 기법은 다목적 최적화 알고리즘의 비수렴성과 앙상블기반기법의 한계점을 극복하였다. 저류층의 생산추이를 입력값, 저류층의 물성을 출력값으로 한 역산모델 신경망을 구성하고, 거리기반 군집화를 통해 실제 생산자료와 유사한 추이를 보이는 후보 저류층을 새로운 학습으로 인공신경망의 지도효과를 개선하였다. 또한, 심층신경망 기법인 오토인코더를 이용하여 학습자료를 내재화함으로써 신경망의 학습속도와 학습성능을 향상시켰다.

제안한 기법의 방법론을 검증하기 위하여 목적함수가 가스 생산량 하나이며, 지질학적 복잡성이 거의 없는 균질한 셰일저류층에서 히스토리매칭을 수행하였다. 제안한 거리기반 후보군 선정 기법을 통하여 새로운 학습집단을 선정함으로써 신경망 기법의 학습성능이 개선됨을 확인하였다. 저류층의 암상분포가 생산량에 큰 영향을 주며, 정규분포를 따르지 않는 채널저류층에서 암상분포를 추정하였다. 앙상블칼만필터는 기본가정인 정규분포를 만족시키지 않는 채널저류층에서 히스토리매칭에 실패한 반면에, 제안한 기법은 암상분포를 신뢰성 있게 추정하였다. 필드 적용성을 평가하기 위하여

불균질 필드에서 오일 생산량, 정저압력, 가스-오일비 등 총 18개의 목적함수에 대한 다목적 히스토리매칭을 수행하였다. 다목적 최적화 알고리즘은 목적함수가 많아 최적해가 수렴하지 못하는 문제가 발생한 반면에, 제안한 기법은 역산모델 신경망을 통하여 목적함수의존적 문제를 극복하여 정확한 예측치를 산출하였다.

기존 기법과의 비교를 통하여 제안한 기법이 거리기반 후보군 선정을 통해 심층신경망의 학습성능을 개선함으로써 신뢰성 있는 최적해를 도출함을 확인하였다. 제안한 기법은 다수의 생산정이 존재하는 저류층에서도 신뢰도 높은 생산량 예측이 가능하므로 저류층 개발을 위한 의사결정의 도구로서 활용될 수 있다.

주 요 어: 거리기반 군집화, 심층신경망, 오테인코더, 히스토리매칭, 다목적 최적화, 저류층 특성화

학 번: 2012-24169

목 차

초 록	I
목 차	III
List of Tables	IV
List of Figures	VI
1. 서론	1
2. 이론적 배경	7
2.1 거리기반 군집화	7
2.2 심층신경망	12
3. 연구 방법	17
3.1 심층신경망 구성	20
3.2 거리기반 후보군을 통한 학습성능 개선	26
4. 개발기법의 검증 및 비교분석 방법	30
4.1 다목적 히스토리매칭 성능분석	31
4.2 현장 적용성 평가	77
5. 결론	98
참고문헌	101
Appendices	108
Abstract	118

List of Tables

Table 3.1 Time steps of PUNQ-S3 field.....	21
Table 4.1 Reservoir properties of the shale reservoir	34
Table 4.2 Neuron description of neural network for shale reservoir....	36
Table 4.3 Errors of initial and trained models	40
Table 4.4 Errors of history matching results	45
Table 4.5 Parameters for training image	47
Table 4.6 Parameters for SNESim simulation.....	47
Table 4.7 Well information and boundary conditions.....	49
Table 4.8 Reservoir properties of channelized reservoir.....	50
Table 4.9 Training error and training time depending on the number of hidden neurons for channelized reservoir.....	52
Table 4.10 Neuron description of autoencoder for channelized reservoir	53
Table 4.11 Neuron description of neural network for channelized reservoir	53
Table 4.12 Errors of history matching results for case 1 (during matching period, 720 days).....	62
Table 4.13 Errors of history matching results for case 1 (during prediction period, 1,800 days).....	62
Table 4.14 Errors of history matching results for case 2 (during matching	

period, 720 days).....	74
Table 4.15 Errors of history matching results for case 2 (during prediction period, 1,800 days).....	74
Table 4.16 Permeability and porosity range for the layers in PUNQ-S3 field.....	78
Table 4.17 Training error and training time depending on the number of hidden neurons for PUNQ-S3 field.....	83
Table 4.18 Neuron description of autoencoder for PUNQ-S3 field.....	84
Table 4.19 Neuron description of neural network for PUNQ-S3 field	84
Table 4.20 Errors of history matching results (oil production rate).....	93
Table 4.21 Errors of history matching results (bottom-hole pressure).	93
Table 4.22 Errors of history matching results (gas-oil ratio)	94
Table 4.23 Errors of history matching results (total oil and gas production)	94

List of Figures

Figure 2.1 The procedures of k-means clustering: (a) step 1, (b) step 2, (c) step 3, (d) step 4.	9
Figure 2.2 Comparison of two clustering methods: (a) k-means clustering, (b) k-medoids clustering.	11
Figure 2.3 Biological neuron(Maltarollo et al., 2013).....	14
Figure 2.4 Schematic of neural network.	14
Figure 2.5 Schematic of autoencoder.	16
Figure 3.1 Workflow of the proposed methodology.	19
Figure 3.2 Autoencoder for channelized reservoir.....	23
Figure 3.3 Autoencoder for PUNQ-S3 field.....	23
Figure 3.4 Schematic diagram of neural network using for proxy model and for inverse model.....	25
Figure 3.5 Distance based selected models for proposed method.	27
Figure 3.6 Distance-based candidate selection for proposed method: (a) channelized reservoir, (b) PUNQ-S3 field.....	29
Figure 4.1 Schematic of shale reservoir with 4 hydraulic fractures.....	33
Figure 4.2 5-years gas production of the shale reservoir.	33
Figure 4.3 Production profile of initial models: (a) daily gas rate, (b) cumulative volume.....	37
Figure 4.4 ANN training for the shale reservoir.	37
Figure 4.5 Production profile of trained models: (a) models trained using	

initial models, (b) the selection of cluster involving the true dataset,	
(c) selected models, (d) models trained using selected models....	38
Figure 4.6 Production profiles of trained models: (a) 1 st generation, (b)	
5 th generation, (c) 10 th generation, (d) 20 th generation.....	39
Figure 4.7 Cumulative volume of trained models: (a) 1 st generation, (b)	
5 th generation, (c) 10 th generation, (d) 20 th generation.....	40
Figure 4.8 Comparison of history matching results for shale reservoir	
(production profile): (a) initial models, (b) Genetic Algorithm, (c)	
Ensemble Kalman-filter, (d) proposed method.....	43
Figure 4.9 Comparison of history matching results for shale reservoir	
(cumulative volume): (a) initial models, (b) Genetic Algorithm, (c)	
Ensemble Kalman-filter, (d) proposed method.....	44
Figure 4.10 Box plot of history matching results.	45
Figure 4.11 Reference field: (a) training image, (b) field.	47
Figure 4.12 Initial models for case 2: (a) training image (orientation: 45),	
(b) field (orientation: 45), (c) training image (orientation: 90), (d)	
field (orientation: 90), (e) training image (orientation: 135), (f) field	
(orientation: 135).	48
Figure 4.13 Permeability distribution of reference field.....	49
Figure 4.14 Training error and training time depending on the number of	
hidden neurons for channelized reservoir.	52
Figure 4.15 Oil production profiles of initial models until 720 days	
(case 1): (a) PROD1, (b) PROD2, (c) PROD3, (d) PROD4.	56

Figure 4.16 Permeability distribution of best fit model (case 1): (a) reference field, (b) initial models, (c) NSGA-II, (d) Ensemble Kalman-filter, (e) proposed method.	57
Figure 4.17 Average error map of permeability (case 1): (a) initial models, (b) matched models by NSGA-II, (c) assimilated models by Ensemble Kalman-filter, (d) selected models by proposed method.	58
Figure 4.18 Production profiles of PROD3 (case 1): (a) (b) NSGA-II (oil, water production), (c) (d) Ensemble Kalman-filter (oil, water production), (e) (f) proposed method (oil, water production).	59
Figure 4.19 Production profiles of PROD2 (case 1): (a) (b) NSGA-II (oil, water production), (c) (d) Ensemble Kalman-filter (oil, water production), (e) (f) proposed method (oil, water production).	60
Figure 4.20 Total production (case 1): (a) (b) NSGA-II (total oil, water production), (c) (d) Ensemble Kalman-filter (total oil, water production), (e) (f) proposed method (total oil, water production).	63
Figure 4.21 Box plot of history matching results at 1,800 days (case 1): (a) Total oil production, (b) Total water production.	64
Figure 4.22 Oil production profiles of initial models until 720 days...	67
Figure 4.23 Permeability distribution of best fit model (case 2): (a) reference field, (b) initial models, (c) NSGA-II, (d) Ensemble Kalman-filter, (e) proposed method.	68

Figure 4.24 Average error map of permeability (case 2): (a) initial models, (b) matched models by NSGA-II, (c) assimilated models by Ensemble Kalman-filter, (d) selected models by proposed method.	69
Figure 4.25 Production profiles of PROD3 (case 2): (a) (b) NSGA-II (oil, water production), (c) (d) Ensemble Kalman-filter (oil, water production), (e) (f) proposed method (oil, water production).	70
Figure 4.26 Production profiles of PROD2 (case 2): (a) (b) NSGA-II (oil, water production), (c) (d) Ensemble Kalman-filter (oil, water production), (e) (f) proposed method (oil, water production).	71
Figure 4.27 Total production (case 2): (a) (b) NSGA-II (total oil, water production), (c) (d) Ensemble Kalman-filter (total oil, water production), (e) (f) proposed method (total oil, water production).	75
Figure 4.28 Box plot of history matching results at 1,800 days (case 2): (a) Total oil production, (b) Total water production.	76
Figure 4.29 Log-permeability distribution of PUNQ-S3: (a) layer 1, (b) layer 2, (c) layer 3, (d) layer 4, (e) layer 5.	79
Figure 4.30 Comparison between production profiles of initial models and reference for PUNQ-S3 field: (a) oil production rate, (b) bottom-hole pressure.	80
Figure 4.31 Comparison between production profiles of initial models and reference for PUNQ-S3 field: (a) gas-oil ratio, (b) watercut.	81

Figure 4.32 Training error and training time depending on the number of hidden neurons for PUNQ-S3 field.	83
Figure 4.33 Log permeability distribution of best fit model for layer 4: (a) reference field, (b) initial models, (c) NSGA-II, (d) Ensemble Kalman-filter, (e) proposed method.	87
Figure 4.34 Average error map of log-permeability for layer 4: (a) initial models, (b) NSGA-II, (c) Ensemble Kalman-filter, (d) proposed method.	88
Figure 4.35 Oil production rate of PROD2: (a) initial models, (b) NSGA-II, (c) Ensemble Kalman-filter, (d) proposed method.	89
Figure 4.36 Bottom-hole pressure of PROD2: (a) initial models, (b) NSGA-II, (c) Ensemble Kalman-filter, (d) proposed method.	90
Figure 4.37 Gas-oil ratio of PROD2: (a) initial models, (b) NSGA-II, (c) Ensemble Kalman-filter, (d) proposed method.	91
Figure 4.38 Total oil production of history matching results: (a) initial models, (b) NSGA-II, (c) Ensemble Kalman-filter, (d) proposed method.	95
Figure 4.39 Total gas production of history matching results: (a) initial models, (b) NSGA-II, (c) Ensemble Kalman-filter, (d) proposed method.	96
Figure 4.40 Box plot of history matching results at 6,025 days: (a) total oil production, (b) total gas production.	97

1 서론

석유개발 현장에서 획득할 수 있는 코어 자료나 탄성과 탐사자료, 물리검층 결과와 같은 저류층의 정적자료로부터 얻을 수 있는 정보는 극히 제한적이며, 지하에 부존하는 석유자원의 특성으로 인하여 저류층의 물성과 유체 유동에 대한 불확실성이 크다. 따라서 신뢰성 있는 저류층 평가를 위해서는 생산자료나 압력자료 등 시간에 따라 얻어진 동적자료를 이용하여 불확실한 저류층의 물성을 예측하는 역산과정이 필요하다. 동적자료와 정적자료를 결합한 역산과정을 통하여 저류층을 특성화하고, 미래의 생산 거동을 예측하는 일련의 과정을 히스토리매칭(history matching)이라 한다.

저류층의 히스토리매칭에 사용되는 대표적인 최적화 기법에는 유전 알고리즘(Genetic Algorithm)이 있다. Holland(1975)는 자연세계의 진화과정을 이용한 최적화 기법인 유전 알고리즘을 제안하였다. 유전 알고리즘은 공간탐색능력이 우수하고, 목적함수를 통하여 우수한 해집합을 효율적으로 산출할 수 있어서 저류층 히스토리매칭을 위한 최적화 기법으로 활발하게 연구되고 있다. Ballester와 Carter(2007)는 실수를 이용한 유전 알고리즘(real-coded genetic algorithm)을 제안하였고, 이를 이용하여 히스토리매칭을 수행하고 미래 생산량을 예측하였다. 다수의 생산정이 있는 저류층에서 유전 알고리즘을 통하여 히스토리매칭을 수행할 때에는 주로 오차 선형가중합법을 이용하여 각각의 목적함수에 가중치를 적용하여 하나의 목적함수로 치환한다(Han et al., 2011). 그런데, 선형가중합법은 최종해 결과가 가중치에 매우 민감할 뿐만 아니라, 각각의 목적함수를 만족시키지 못하는 해가 도출되기도 하는 한계가 있다.

이러한 단목적 유전 알고리즘의 한계를 개선하기 위하여 Deb 등(2002)은 비지배 정렬기반의 유전 알고리즘(Non-dominated Sorting Genetic Algorithm-II, NSGA-II)을 제안하였다. NSGA-II는 비지배 정렬을 통해 다수의 목적함수에 대해서 서로 지배관계에 속하지 않은 해들을 선택할 수 있기 때문에 가중치에 의한 영향을 받지 않으면서 각각의 목적함수를 만족시키는 최종해를 도출할 수 있다는 장점이 있다. Kim 등(2017)은 셰일저류층에서 두 가지 목적함수(가스 생산량, 누적 가스생산량)에 대하여 NSGA-II를 이용하여 히스토리매칭을 수행하였고, 단목적 유전 알고리즘을 통한 히스토리매칭 결과와 비교하였다. 비교 결과, NSGA-II가 단목적 유전 알고리즘에 비하여 해의 다양성을 보존하면서도 두 가지 목적함수를 모두 만족시키는 최적해를 도출함을 확인하였다. 그러나 NSGA-II는 목적함수가 많아지면 최적해가 수렴하지 못하는 목적함수 의존적 문제가 존재한다.

생물 내에서 이루어지는 신경세포의 전달과정에서 영감을 얻은 인공신경망은 학습(training)을 통하여 입력층과 출력층 사이의 상관성을 확보하는 기술이다. 학습에 기반한 인공신경망 모델은 정방향 모델링에 기반한 프록시 모델(proxy model)로써 저류층 시뮬레이션을 대체하여 왔다(Park et al., 2005; Min et al., 2011). Siripatrachai와 Ertekin(2012)은 셰일가스 저류층에서 인공신경망을 이용한 생산량 예측모델을 구축하였다. 구축한 인공신경망을 이용하여 저류층 시뮬레이션을 수행하지 않고도 새로운 셰일가스 저류층의 생산량을 산출하였으나, 10% 이상의 오차를 보이는 한계를 보였다. Kim 등(2015)은 인공신경망의 예측성능을 높이기 위하여 민감도분석 결과를 반영한 셰일가스 저류층 생산량 예측 인공신경망을 구축하였다. Shahkarami 등(2014)은 유정의 생산자료를

입력층, 저류층의 물성을 출력층으로 하는 인공신경망을 제안하여 인공신경망이 석유공학분야에서 역산 모델로써 활용 가능성을 보였다. 안성인(2017)은 2차원 채널저류층에서 인공신경망을 이용한 저류층 특성화 기법을 적용하였다. 심층신경망(Deep Neural Network, DNN) 중 하나인 오토인코더(auto encoder) 기법을 적용한 저류층 특성화 결과와 일반적인 인공신경망을 적용한 특성화 결과를 비교하였다. 오토인코더 기법이 일반적인 인공신경망과 비교하여 채널저류층의 채널 연결성을 보존하면서도 생산량 예측결과의 신뢰성이 높은 것을 확인하였다. 하지만, 한번의 역산을 통하여 특성화를 완료하였기 때문에 인공신경망의 학습성능을 개선하는 방법은 고려하지 못하였다는 한계가 있다.

거리기반 군집화는 모델간의 차이를 거리로 정의한 거리기반 지도상에서 가까운 모델끼리 분류함으로써 유사한 특징을 갖는 모델들을 군집화하는 기법이다. 대표적인 거리기반 군집화 기법으로는 k-평균 군집화(k-means clustering)와 k-중간점 군집화(k-medoids clustering)가 있다(Berkhin, 2006). k-평균 군집화는 계산시간이 빠르다는 장점이 있지만, 이상점(outlier)이 있을 경우 제대로 된 군집화가 이루어지지 못한다는 단점이 있다(Velmurugan, et al., 2010). k-중간점 군집화는 k-평균 군집화와 비하여 계산시간이 느리지만, 이상점에 취약한 k-평균 군집화의 단점을 보완한 기법이다(Kaufman and Rousseeuw, 1987). 저류층 물성 등의 정적자료를 거리로 정의하여 거리기반 군집화를 수행하면 대표모델에 대해서만 저류층 시뮬레이션을 수행함으로써 시뮬레이션 계산비용을 감소시킬 수 있다. 또한, 저류층의 생산량 등의 동적자료를 거리로 정의하여 대상 저류층의 생산이력과 같은 군집에 속하는 저류층 모델을 찾는 과정을 통하여 좋은 형질의 저류층을 선택할 수 있다.

앙상블기반기법(ensemble-based method)은 저류층 특성화와 저류층의 불확실성 평가에 사용되고 있는 방법이다. Naevdal과 Vefring(2002)은 앙상블칼만필터(Ensemble Kalman-filter, EnKf)를 석유공학 분야에 처음으로 적용하여 사암저류층에 대한 저류층 특성화를 수행하였다. 앙상블칼만필터를 통하여 유체투과도를 교정하고, 생산량 예측과 불확실성 평가를 수행하여 신뢰성 있는 결과를 산출하였다. 임서진(2016)은 2차원 채널저류층에서 다수의 트레이닝 이미지(training image)를 통하여 초기 저류층 모델을 생성한 뒤, k-중간점 군집화를 통하여 참조필드의 생산이력이 유사한 군집을 선정하였다. 해당 군집의 저류층 모델들을 생성한 트레이닝 이미지를 역추적하여 트레이닝 이미지의 확률을 계산하였고, 계산한 확률을 이용하여 새롭게 초기 저류층 모델을 생성하였다. Lee 등(2013)은 채널저류층에서 거리기반 앙상블스무더(Ensemble Smoother, ES)를 이용한 저류층 특성화 기법을 제안하였다. 트레이닝 이미지와 코어 자료를 이용하여 정적자료 기반의 다수의 저류층 모델을 생성하고, 생성된 저류층 모델들 간의 비유사성을 거리로 정의하여 거리기반 군집화를 수행하였다. 거리기반 군집화를 통해 분류된 군집마다 다른 칼만게인을 적용하여 거리기반 앙상블스무더를 수행한 결과, 기존의 앙상블스무더나 앙상블칼만필터와 비교하여 신뢰성 있는 결과를 산출하였다.

앙상블기반기법의 두 가지 핵심가정은 앙상블의 개수가 충분히 크며, 관측오차가 정규분포를 따른다는 가정과 앙상블의 평균을 참값으로 가정하는 것이다. 앙상블기반기법은 이 두 가지 핵심가정을 만족하는 경우에서 신뢰성 있는 히스토리매칭 결과를 보여준다(Gu and Oliver, 2006; Evensen et al., 2007). 반면에, 핵심가정을 만족하지 못하는 경우에는 오버슈팅(over shooting)과 필터발산(filter

divergence)과 같은 문제가 발생한다는 한계점이 있다(Aanonsen, et al., 2009; Oliver and Chen, 2011). 오버슈팅은 모델의 특성값이 지나치게 크거나 작은 값으로 교정되어 물리적으로 불가능한 값이 발생하는 문제로써 앙상블기반기법이 교정하는 특성값의 물리적인 의미를 고려하지 못하기 때문에 생기는 한계이다. 필터발산은 다양한 특성값을 가지는 초기모델들이 교정을 통하여 하나의 저류층 모델로 수렴하게 되는 문제이다. 필터발산 문제가 생기면 모델들끼리 유사성이 매우 높기 때문에 추정오차공분산이 작아져 교정이 제대로 이루어지지 못하며, 불확실성평가도 불가능해진다.

이상의 저류층 히스토리매칭을 수행한 기존의 연구에는 다음과 같은 한계가 있다.

- (1) 진화 기반의 최적화 기법인 유전 알고리즘은 최적화 인자의 개수가 증가함에 따라서 계산비용이 크게 증가한다는 단점이 있다. 또한, 단목적 유전 알고리즘은 최적해를 찾아가는 과정에서 하나의 해로 수렴하는 경향이 있기 때문에 해의 다양성을 보존하지 못한다는 한계가 있다.
- (2) 비지배 정렬 기반의 다목적 유전 알고리즘인 NSGA-II는 단목적 유전 알고리즘과 달리 가중치에 의한 영향을 받지 않는 방법으로 2개 이상의 목적함수에 대해 신뢰성 있는 최적화 결과를 보여준다. 하지만, 목적함수의 개수가 증가하면 최적해들이 수렴하기 어렵다는 한계가 있다.

(3) 인공신경망 기법을 석유공학에 적용한 대부분의 연구는 인공신경망을 저류층 시뮬레이션을 대신할 수 있는 프록시 모델로써 활용하였다. 인공신경망을 역산에 활용하여 2차원 채널저류층 히스토리매칭을 수행한 연구에서는 한번의 역산으로 히스토리매칭을 종료하였다. 따라서 인공신경망 학습성능에 대한 개선이 이루어지지 못했다는 한계가 있다.

(4) 앙상블기반기법은 관측오차가 정규분포를 따른다고 가정하고, 앙상블의 평균을 참값으로 가정한다. 이와 같은 가정들을 만족하지 않는 저류층에 앙상블기반 기법을 적용하면, 오버슈팅과 필터발산과 같은 문제점이 발생한다는 한계가 있다.

이 연구의 핵심목적은 다수의 목적함수를 동시에 만족시키는 최적해를 찾는 최적화 기법 개발을 통한 생산량 예측 및 불확실성평가이다. 생산자료를 효과적으로 히스토리매칭하는 심층신경망 기반의 역산모델을 개발하고, 거리기반 후보군 선정을 통해 우수한 학습표본집단을 선정함으로써 인공신경망의 학습성능을 개선할 수 있는 방법론을 제안하였다. 제안한 기법은 다수의 생산정이 존재하는 현장의 저류층에서 정확한 생산량 예측이 가능하여 향후 저류층 운영계획 수립을 위한 핵심적 도구로 활용될 것으로 기대된다.

2 이론적 배경

이 연구에서는 거리기반 군집화기법과 비선형 관계에 대한 신경망 학습을 통하여 다목적 히스토리매칭을 수행하는 기법을 개발하였다. 이 장에서는 제안한 기법에서 사용한 이론적 배경에 대해 소개한다. 2.1절에서는 이 연구에서 거리기반 후보군 선정을 할 때 사용한 거리기반 군집화 기법에 대하여 설명한다. 2.2절에서는 신경망학습을 통하여 새로운 모델을 매칭할 때 사용한 심층신경망에 대하여 소개한다.

2.1 거리기반 군집화

(1) k-평균 군집화

k-평균 군집화는 주어진 데이터를 설정한 개수의 군집으로 묶는 알고리즘이다. 이 때, 각 군집의 평균과 군집 내의 다른 데이터와의 거리의 제곱 합을 비용 함수(cost function)으로 정의하고, 이 비용 함수 값을 최소화하는 방향으로 군집을 새롭게 설정하면서 군집화를 수행한다(Hartigan and Wong, 1979).

식 (2.1)은 k-평균 군집화에서 정의하는 비용 함수를 나타낸다.

$$\text{Cost function} = \sum_{i=1}^k \sum_{x_j \in S_i} |x_j - u_i|^2 \quad (2.1)$$

위 식에서 k 는 군집화하고자 하는 군집의 개수, S_i 는 i 번째

군집을 의미하며, u_i 는 i 번째 군집의 중심이다. x_j 는 S_i 에 속한 임의의 모델을 의미하며, m 은 S_i 에 속한 모델의 개수이다. 이 비용함수의 값이 최소가 되게 하는 군집 S_i 를 찾는 것이 k -평균 군집화의 목표이다. k -평균 군집화는 다음과 같은 과정을 거쳐서 군집화를 수행하게 된다.

(1) Figure 2.1(a)와 같이 전체 n 개의 모델 중에서 임의의 모델 k 개를 초기 평균값으로 선정한다($k < n$, Figure 2.1에서 $k = 3$).

(2) 모든 모델들을 거리상으로 가장 가까이에 있는 평균값과 같은 군집에 속하게 설정한다(Figure 2.1(b))

(3) 각 군집에 속한 모델들의 평균값을 계산하고, 이 값을 새로운 평균값으로 재조정한다(Figure 2.1(c)).

(4) 각 군집의 평균값이 변하지 않을 때까지 (2)~(3) 과정을 반복한다(Figure 2.1(d)).

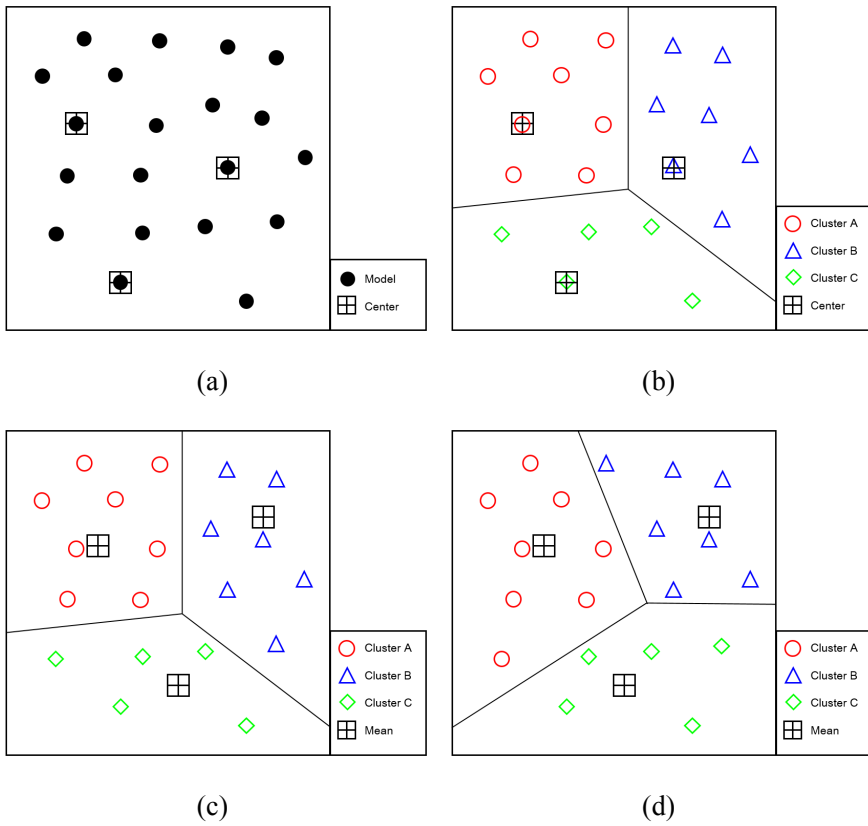


Figure 2.1 The procedures of k-means clustering:
 (a) step 1, (b) step 2, (c) step 3, (d) step 4.

(2) k -중간점 군집화

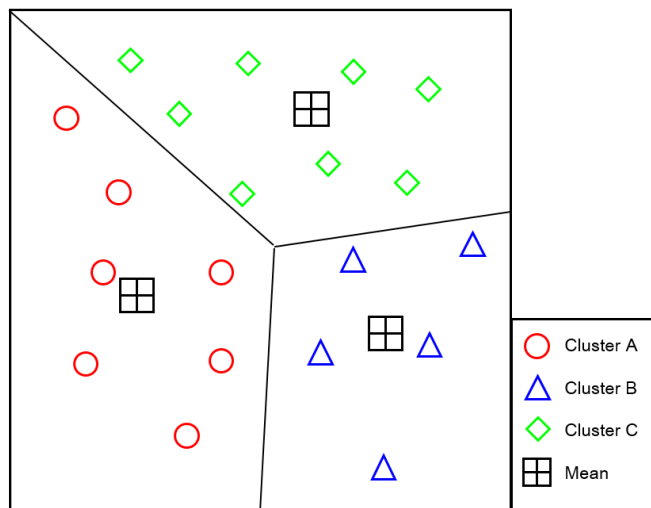
k -중간점 군집화는 Kaufman과 Rousseeuw(1987)가 제안한 군집화 알고리즘으로써 이상점에 취약한 k -평균 군집화 알고리즘의 단점을 보완하였다. k -평균 군집화에서는 각 군집의 평균값을 군집의 중심으로 설정하는 반면에, k -중간점 군집화에서는 각 군집의 평균값에 가장 가까운 모델을 군집의 중간점으로 설정한다(Figure 2.2). k -중간점 군집화를 수행하는 과정은 다음과 같다.

(1) 전체 n 개의 모델 중에서 임의의 모델 k 개를 초기 평균값으로 선정한다. k 는 군집의 개수이다($k < n$).

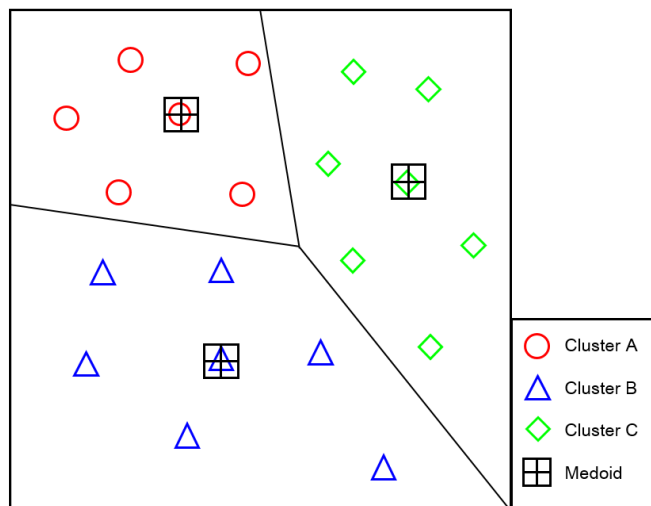
(2) 모든 모델들을 거리상으로 가장 가까이에 있는 중간점과 같은 군집에 속하게 설정한다.

(3) 각 군집에 속한 모델들의 평균값을 계산하고, 이 값과 가장 가까운 모델을 새로운 중간점으로 재조정한다.

(4) 각 군집의 중간점이 변하지 않을 때까지 (2)~(3) 과정을 반복한다.



(a)



(b)

Figure 2.2 Comparison of two clustering methods:
(a) k-means clustering, (b) k-medoids clustering.

2.2 심층신경망

(1) 인공신경망

인공신경망은 생물학적 신경세포에서 영감을 얻은 통계학적 학습 알고리즘이다. 생물학적 신경세포는 수상돌기(dendrites), 세포체(cell body), 축삭돌기(axon), 축삭종말(terminal axon)로 구성된다(Figure 2.3). 이전 신경세포에서 보내는 전기신호는 시냅스를 통해 다음 신경세포의 수상 돌기로 전달된다. 수상 돌기는 이 전기 신호를 받아들여 세포체에 전달하는 역할을 한다. 세포체에 전달된 신호는 다시 축삭돌기와 축삭종말을 통해 다음 신경세포에게 전기신호로 전달된다.

인공신경망은 생물학적 신경세포의 신호전달과정을 모방하여 비선형적 관계를 가진 입력정보와 출력정보에 대한 처리를 수행한다. 기본적으로 인공신경망은 Figure 2.4와 같이 한 개의 입력층(input layer), 한 개의 출력층(output layer), 한 개 이상의 은닉층(hidden layer)로 구성된다. 각각의 층에는 여러 개의 신경세포가 있고, 신경세포 내부에서는 가중치의 합과 같은 단순한 계산만을 수행한다. 각 층의 신경세포와 다음 층의 신경세포를 연결하는 것이 시냅스이며, 시냅스들은 두 신경세포 간의 연결 강도를 의미하는 가중치 값을 저장한다. 인공신경망을 정의하는 세가지 인자는 다음과 같다.

- 1) 다른 층의 신경세포간의 연결패턴
- 2) 연결의 가중치를 갱신하는 학습
- 3) 가중치를 이용하여 출력하는 활성화 함수

신경세포간의 연결패턴은 신호가 전달되는 방향에 따라서 구분할

수 있다. 대표적인 방법으로는 전방 전달 신경망(feed forward networks)과 순환 신경망(recurrent neural network)이 있다. 전방 전달 신경망은 신호가 입력층에서 출력층으로 가는 단방향으로만 전달되는 신경망이다. 순환 신경망은 전방 전달 신경망과는 다르게 신경세포간의 신호가 양방향으로 전달이 가능한 신경망이다.

인공신경망의 학습은 인공신경망을 통해 예측한 예측값과 실제 값 사이의 예측오차를 최소화 하는 비선형 함수를 찾아내기 위하여 가중치를 조절하는 과정이다. 학습 알고리즘으로는 다양한 최적화 알고리즘이 사용되며, 인공신경망에서 대표적으로 사용되는 학습 알고리즘은 역전파(back propagation) 기법을 이용하여 실제 기울기를 계산하는 경사 하강법(gradient descent)이다. 경사 하강법은 함수를 미분하여 기울기(경사)를 구한 다음에, 기울기가 낮은 방향으로 조금씩 인자를 바꾸면서 최적해를 구하는 방법이다(Qian, 1999). 신경세포 내부에서 이루어지는 가중치의 합은 연산자를 통하여 계산하며, 여기에서 계산된 값을 최종적으로 활성화 함수를 이용하여 구현한다. 인공신경망에서 사용되는 활성화 함수에는 S자형(sigmoid), 선형(Liner), 문턱형(Threshold) 등이 있다.

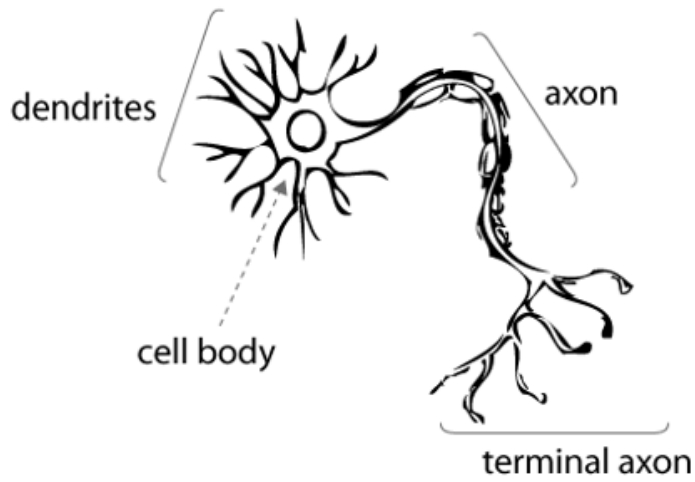


Figure 2.3 Biological neuron(Maltarollo et al., 2013).

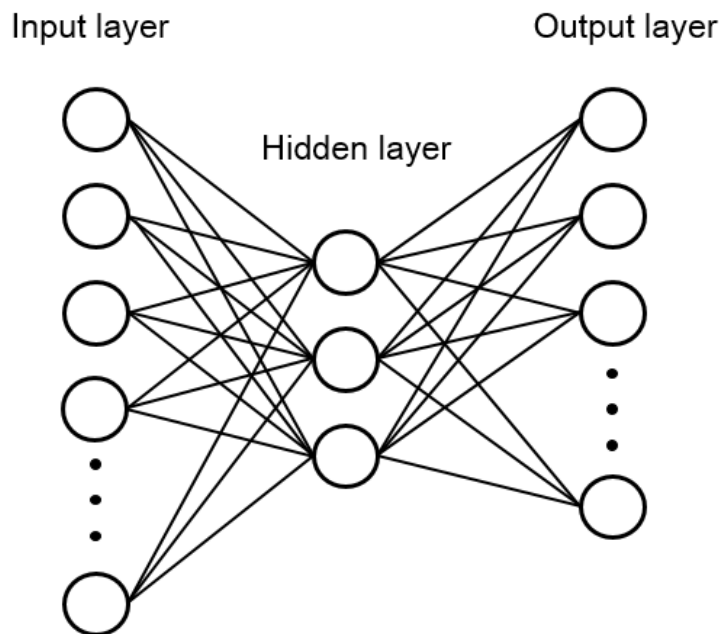


Figure 2.4 Schematic of neural network.

(2) 심층신경망

심층신경망은 입력층과 출력층 사이에 여러 개의 은닉층들로 이뤄진 신경망이다(Schmidhuber, 2015). 심층신경망은 일반적인 인공신경망과 마찬가지로 복잡한 비선형 관계를 모사할 수 있는 방법으로써, 인공신경망에 비하여 더 많은 추가 계층들을 사용하기 때문에 하위 계층들의 특징을 파악할 수 있다. 심층신경망은 많은 수의 은닉층을 활용하여 정보를 처리하기 때문에 연산비용이 증가한다는 단점이 있지만, 최근 컴퓨터의 연산처리능력이 발전함에 따라 그래픽 모델링, 패턴인식, 신호처리, 최적화 등 다양한 분야에서 활용되고 있다(Li and Yu, 2014).

일반적인 인공신경망 기법에서 많은 수의 은닉층을 사용하게 되면 가중치가 거의 변하지 않아서 학습 성능이 제대로 개선되지 못하는 문제가 발생한다. 이러한 문제점을 극복하기 위하여 각 층의 신경세포들을 부호화(encoding)하는 오토인코더(autoencoder)가 제안되었다(Hinton et al., 2006; Bengio et al., 2007). 심층신경망 기법 중 하나인 오토인코더는 자료의 압축코드화, 이미지 변환, 복구 등 많은 분야에 활용되고 있다. 오토인코더는 Figure 2.5와 같이 입력층, 3개의 은닉층, 출력층으로 구성된다. 두번째 은닉층은 첫번째 은닉층을 이용하여 입력층을 부호화한 층이며, 세번째 은닉층을 이용하여 두번째 은닉층을 재구현(decoding)하면 입력층과 같은 차원의 출력층이 된다. 기본적으로 오토인코더는 입력층의 원본자료를 압축하는 데에 사용되며, 이 때에는 은닉층의 개수가 입력층의 개수보다 작다.

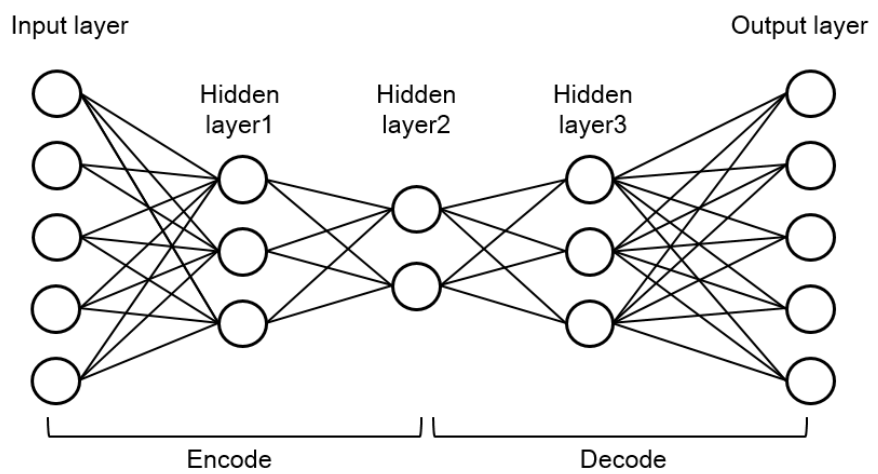


Figure 2.5 Schematic of autoencoder.

3 연구 방법

히스토리매칭의 궁극적인 목표는 신뢰성 있는 저류층 모델을 추정하는 것이다. 해당 저류층에서 확보 가능한 동적자료(오일 생산량, 가스 생산량, 저류층 정저압력 등)와 유사한 동적자료를 나타내는 모델들을 생성함으로써 저류층의 미래 생산량을 예측할 수 있다. 이와 같은 히스토리매칭을 통하여 해당 저류층에 대한 가치평가를 수행할 수 있으며, 향후 저류층 운영계획을 세울 수 있다.

제안한 기법의 순서도는 Figure 3.1과 같으며, 순서도의 각 과정은 다음과 같다.

- (1) 해당 저류층의 저류층 물성정보를 토대로 매칭하고자 하는 저류층 물성의 범위를 산정한 뒤, 다수의 초기 저류층 모델을 생성한다.
- (2) 저류층 시뮬레이션을 이용하여 초기 저류층 모델들의 동적자료를 산정한다.
- (3) 오토인코더를 통하여 초기 저류층 모델들의 저류층 특성을 부호화한다.
- (4) 초기 저류층 모델들의 동적자료를 입력자료, 저류층 물성을 출력자료로 하는 비선형 관계에 대한 신경망학습을 수행한다.
- (5) 신경망학습 결과를 이용하여 대상 저류층의 동적자료와 유사한 동적자료를 나타내는 새로운 저류층 모델들을 생성한다.

- (6) 신경망 학습을 통해 생성된 새로운 저류층 모델들의 동적자료를 산정한다.
- (7) 학습에 사용한 저류층 모델들과 새롭게 생성된 저류층 모델들, 대상 저류층의 동적자료를 거리기반 지도에 나타내어 거리기반 후보군을 선정한다.
- (8) 초기에 설정한 횟수만큼 선정된 거리기반 후보군들을 대상으로 다시 (4)~(7)과정을 반복한다.
- (9) 마지막으로 선정된 거리기반 후보군을 최종 히스토리매칭 결과로 산출한다.

이 장에서는 제안한 기법을 통하여 히스토리매칭을 수행하는 과정에 대해 소개한다. 3.1절에서는 심층신경망을 이용한 학습과정에 대하여 설명한다. 3.2절에서는 거리기반 후보군 선정을 통해 심층신경망의 학습성능을 개선하는 방법에 대하여 설명한다.

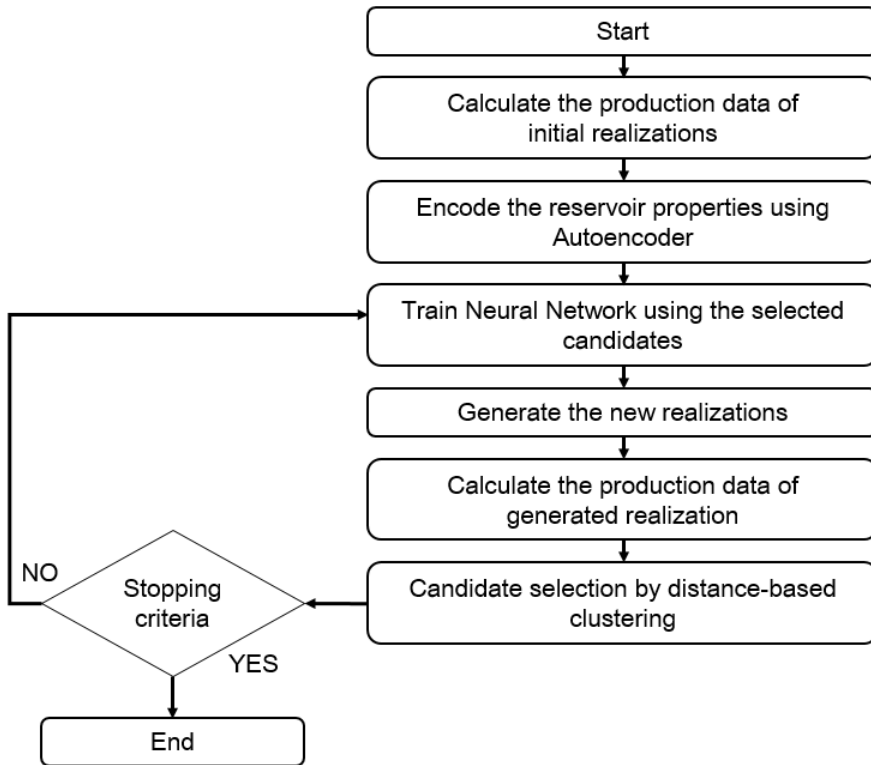


Figure 3.1 Workflow of the proposed methodology.

3.1 심층신경망 구성

(1) 입력자료 표준화

석유공학에서 신경망학습을 활용한 대부분의 연구에서는 신경망 학습결과를 저류층 시뮬레이션을 대신할 수 있는 프록시 모델(proxy model)로 사용하였다. 하지만, 이 연구에서는 신경망 학습결과를 역산 모델(inverse model)로 활용하여 대상 저류층과 유사한 저류층 모델을 생성하였다. 신경망학습에 사용하는 입출력 학습 자료는 수치적 안정성을 위하여 표준화를 해야 한다(Centilmen et al., 1999). 이 연구에서는 식 (3.1)과 같이 각 모델의 생산량과 대상 저류층의 생산량의 차이를 표준화하여 입력자료로 설정하였다. 이와 같은 표준화 과정을 통하여 입력자료를 동일한 범위로 재조정하였다.

$$I_k(t) = 2 * \frac{d_k(t) - d_{min}(t)}{d_{max}(t) - d_{min}(t)} - 1, (where, d_k(t) = |P_k(t) - P_{true}(t)|) \quad (3.1)$$

위 식에서 $I_k(t)$ 는 k 번째 모델의 t 시점에서의 입력자료, $d_k(t)$ 는 k 번째 모델의 t 시점에서의 생산량 차이, $P_k(t)$ 는 k 번째 모델의 t 시점에서의 생산량, $P_{true}(t)$ 는 t 시점에서의 실제 생산이력이다.

세일가스 저류층의 경우에는 가용한 동적자료가 가스 생산량뿐이기 때문에 하나의 가스 생산량만을 뉴런벡터로 설정하였다. 30일 단위의 가스생산량을 이용하여 720일 동안 히스토리매칭을 수행하였기 때문에 세일가스 저류층에서 사용한 입력층의 뉴런수는 총 24개이다. 채널저류층에서는 가용한 동적자료가 4개 생산정의 오일 생산량과 물 생산량, 총 8개이기 때문에 8개의 동적자료를 합친 뉴런벡터를 이용하였다. 세일가스 저류층과 마찬가지로 30일 단위의

오일, 물 생산량을 이용하여 720일 동안 히스토리매칭을 수행하였고, 입력층의 뉴런수는 총 $8 \times 24 = 192$ 개이다. PUNQ-S3 필드의 경우에는 가용한 동적자료가 6개 생산정의 오일 생산량, 정저압력, 가스-오일 비율, 총 18개이기 때문에, 18개의 동적자료를 합친 뉴런벡터를 이용하였다. PUNQ-S3 필드에서는 “shut-in”을 여러 번 하기 때문에 “shut-in”하는 동안에는 생산이 이루어지지 않는다. 따라서 Table 3.1과 같이 Gu와 Oliver(2005)가 사용한 20개의 시점을 사용하였고, 입력층의 뉴런수는 총 $18 \times 20 = 360$ 개이다.

Table 3.1 Time steps of PUNQ-S3 field

Time step	1	2	3	4	5
Time (days)	1.01	91	182	274	366 (shut-in)
Time step	6	7	8	9	10
Time (days)	1461 (shut-in)	1642	1826	1840 (shut-in)	1841
Time step	11	12	13	14	15
Time (days)	2008	2192	2206 (shut-in)	2373	2557
Time step	16	17	18	19	20
Time (days)	2571 (shut-in)	2572	2738	2922	2936 (shut-in)

(2) 오토인코더

세일가스 저류층에서는 7가지의 저류층 물성을 사용하였기 때문에 오토인코더를 이용한 출력자료 압축을 하지 않았다. 채널저류층과 PUNQ-S3 필드에서는 격자별 유체투과도에 대한 히스토리매칭을 수행하였기 때문에 입력자료의 크기보다 출력자료의 크기가 크다. 출력자료의 크기가 입력자료에 비하여 일정 수준 이상 커지게 되면 신경망학습에서 연산속도가 급격하게 느려지고 예측성능도 감소하게 되는 문제가 발생하기 때문에, 오토인코더를 이용하여 격자별 유체투과도를 압축변환 하였다.

Figure 3.2는 채널저류층에서 사용한 오토인코더의 모식도를 나타낸다. 인공신경망 학습을 수행하기에 앞서, 출력자료인 격자별 유체투과도 2,601개를 오토인코더기법을 이용하여 100개의 자료로 압축하였다. Figure 3.3은 PUNQ-S3 필드에서 사용한 오토인코더의 모식도이다. PUNQ-S3 필드에서는 2,660개의 격자별 유체투과도를 100개의 자료로 압축하였다.

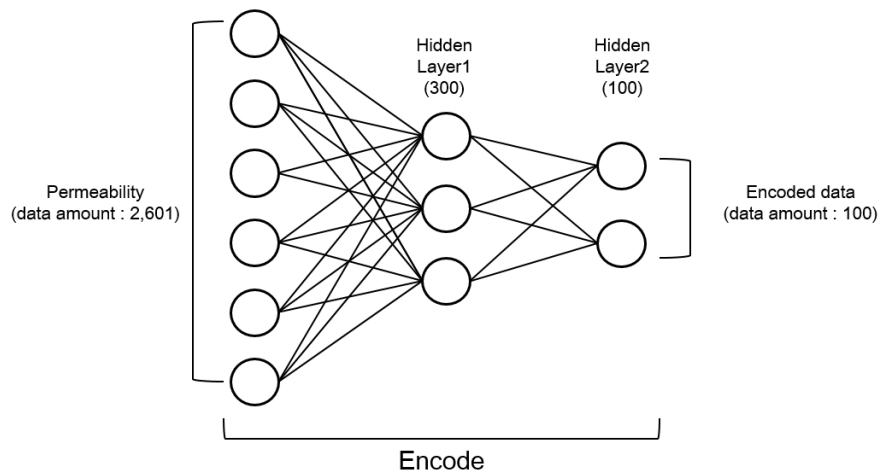


Figure 3.2 Autoencoder for channelized reservoir.

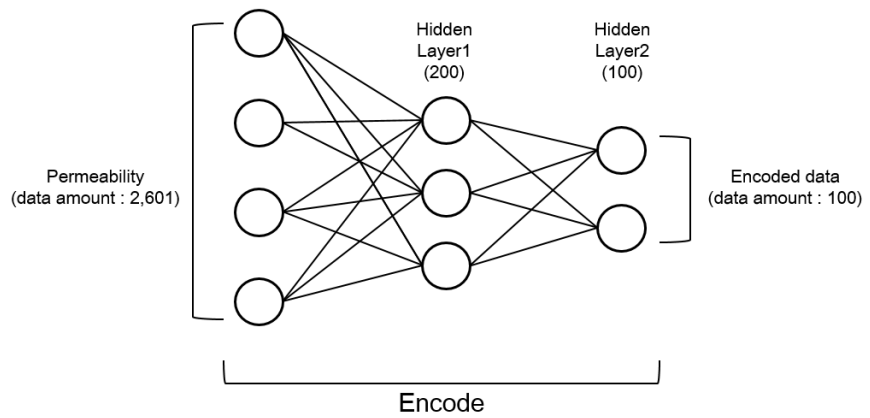


Figure 3.3 Autoencoder for PUNQ-S3 field.

(3) 신경망 학습

신경망 학습에서 사용되는 입출력 자료 쌍은 일반적으로 세 가지로 구분된다. 첫 번째는 학습자료 쌍(training set)으로 자료를 이용하여 가중치를 변화시켜가며 학습을 하기 위한 자료 쌍이다. 두 번째는 검증자료 쌍(validation set)으로 학습이 타당하게 진행되고 있는지를 판별하기 위한 자료 쌍이다. 네트워크의 검증 자료쌍 오차가 최소가 되는 지점에서 학습을 종료 시키며, 이 시점에서의 가중치와 편향 값을 저장하여 사용하게 된다(Araghinejad, 2013). 세 번째는 테스트자료 쌍(test set)으로 학습자료 쌍과 검증자료 쌍으로 사용되지 않으며, 신경망의 성능을 평가비교하기 위해 사용된다. 각각의 자료 쌍들은 학습 가능한 자료 쌍의 수에 따라 분배할 수 있으며, 이 연구에서는 전체 자료 쌍으로부터 학습자료 쌍 80%, 검증자료 쌍 10%, 테스트자료 쌍 10%로 분배하였다.

신경망으로는 피드포워드 신경망을 사용하였다. 피드포워드 신경망은 인공신경망 분야에서도 널리 적용될 뿐만 아니라 다른 신경망보다 빠른 연산이 가능하며, 많은 자료 쌍을 가진 문제를 처리할 수 있는 장점이 있다(Doraisamy et al., 1998). Figure 3.4는 인공신경망을 프록시 모델과 역산 모델로 활용하는 과정을 비교한 모식도이다. 인공신경망을 프록시 모델로 활용하는 연구들에서는 상단에 있는 화살표의 방향대로 좌측의 저류층 물성을 입력층, 우측의 생산이력을 출력층으로 하는 신경망 학습을 수행한다. 반면에, 역산 모델로 활용한 이 연구에서는 그림 하단에 있는 화살표의 방향대로 우측의 생산이력을 입력층, 좌측의 저류층 물성을 출력층으로 하는 신경망 학습을 수행한다.

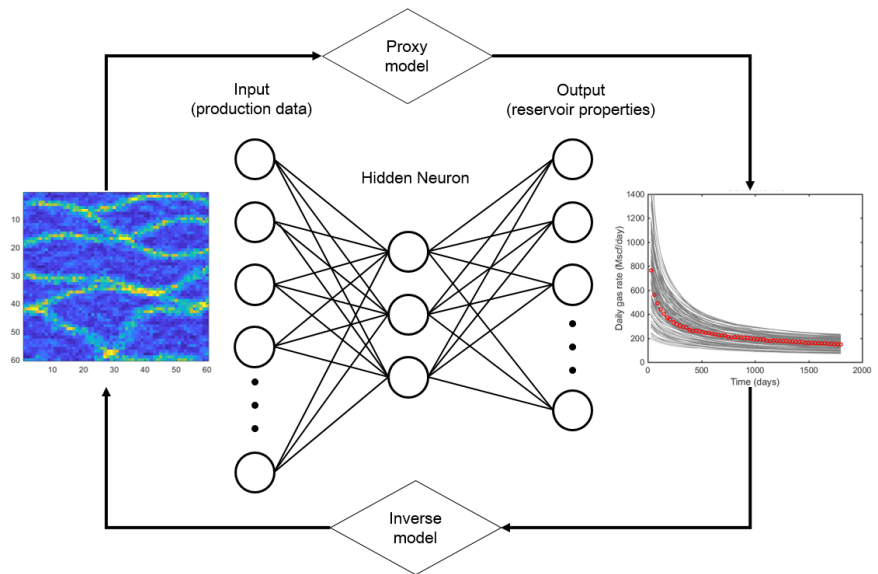


Figure 3.4 Schematic diagram of neural network using for proxy model and for inverse model.

3.2 거리기반 후보군을 통한 학습성능 개선

이 연구에서는 k -중간점 군집화를 이용하여 대상 저류층과 저류층 모델들의 생산량을 거리로 정의하여 군집화하였다. 생산량을 기준으로 수행한 군집화에서 대상 저류층과 같은 군집에 속한 좋은 형질의 저류층 모델들을 거리기반 후보군으로 선정하여 신경망학습의 성능을 개선하였다.

한편, 대상 저류층과 같은 군집에 속한 저류층들만 선정하면, 대상 저류층과 유사한 생산이력을 가지고 있음에도 다음 세대의 학습자료로 선정되지 못하는 경우가 발생할 수 있다. Figure 3.5(a)는 저류층 모델들에 대하여 거리기반 군집화를 수행하기 전의 모습을 보여준다. Figure 3.5(b)는 k -중간점 군집화를 수행한 결과이며, 동그라미는 군집 A, 세모는 군집 B, 다이아몬드는 군집 C, 색칠된 다이아몬드가 대상 저류층 모델을 의미한다. 여기에서 대상 저류층이 속한 군집 C만을 다음 세대의 학습자료로 선정하게 되면, 대상 저류층과 가까이에 있는 군집 B의 모델들을 고려하지 못하게 된다.

Figure 3.5(c)는 대상 저류층을 중심으로 군집 C의 중간점과의 거리를 반지름으로 하는 원을 그린 모습이다. 군집 B 중에서 색칠된 3개의 모델이 원안에 포함되는 것을 알 수 있다. 이처럼 대상 저류층과 가까운 거리에 있음에도 선정되지 않는 문제를 방지하기 위하여, 이 연구에서는 Figure 3.5(d)와 같이 대상 저류층과 같은 군집이 아니더라도 대상 저류층과 가까운 거리에 있는 모델들은 거리기반 후보군으로 추가 선정하였다. 이 때, 거리의 기준은 대상 저류층과 군집의 중간점과의 거리로 설정하였다.

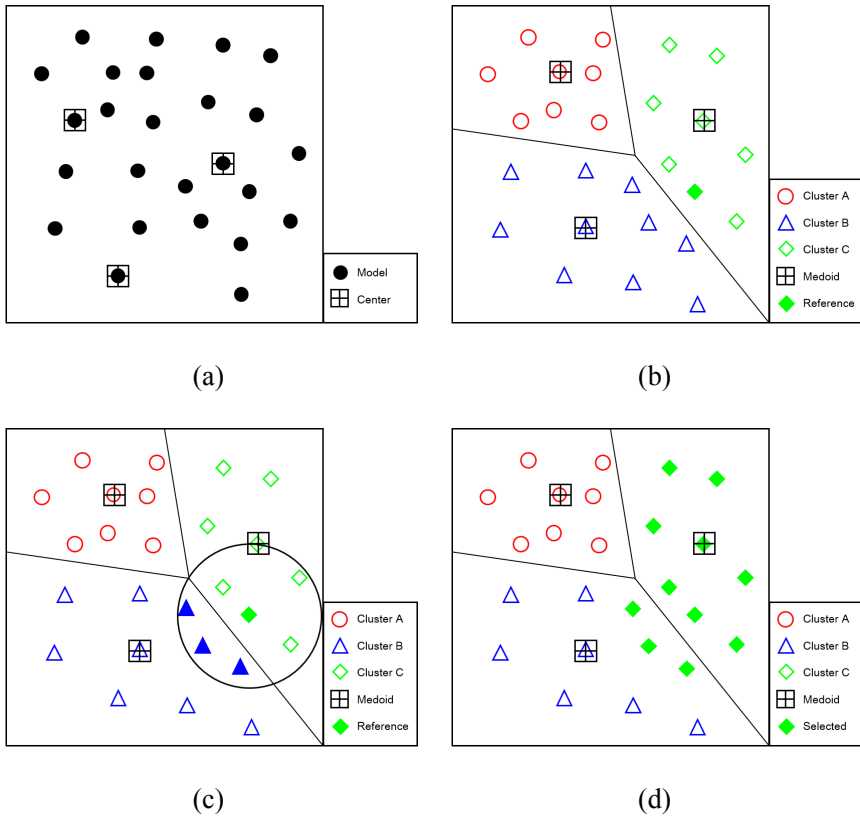


Figure 3.5 Distance based selected models for proposed method.

이 연구에서는 식 (3.2)와 같이 각 저류층 모델들 간의 생산량 차이를 거리로 정의하였다.

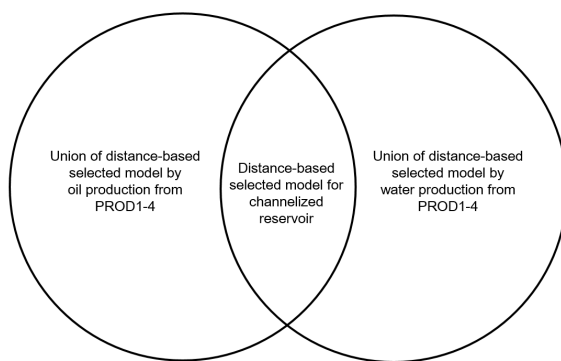
$$d_{ij} = \sum_{t=0}^{N_t} \sqrt{\left| \frac{P_i(t)}{P_{max}(t)} - \frac{P_j(t)}{P_{max}(t)} \right|^2} \quad (3.2)$$

위 식에서 d_{ij} 는 i 번째 모델과 j 번째 모델의 거리, N_t 는 마지막 시점, $P_i(t)$ 는 i 번째 모델의 t 시점에서의 생산량이다. $P_{max}(t)$ 는 t 시점에서 모든 모델의 생산량 중에서 가장 큰 생산량을 의미한다. 이렇게 구한 $N_t \times N_t$ 거리행렬을 다차원축소법(Seber, 1984)을 통하여 2차원으로 축소하여 Figure 3.5(a)와 같이 거리기반지도에 나타내었다.

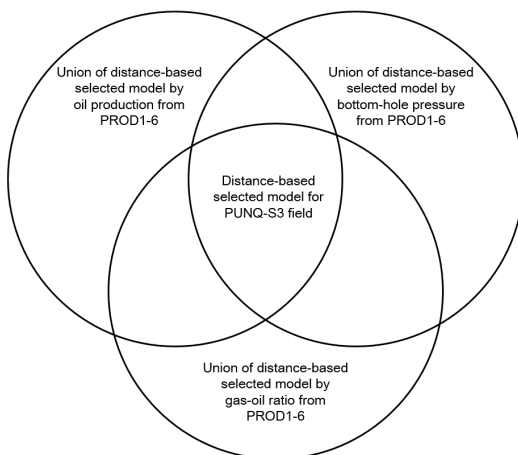
고려할 생산량이 하나인 세일저류층에서는 가스 생산량에 대해서만 거리를 계산하고, 거리기반 후보군을 선정하여 다음 세대 신경망학습의 학습자료로 사용하였다. 채널저류층과 PUNQ-S3 필드에서는 고려해야 할 생산량이 각각 8개, 18개이기 때문에, 모든 생산량에 대하여 같은 방식으로 거리기반 후보군을 선정하였다. 각각의 생산량별 거리기반 후보군들의 합집합은 거의 대부분의 후보군들이 선정되기 때문에 좋은 형질의 모델을 선정할 수가 없고, 반대로 교집합은 매우 작은 수의 모델만 선정되기 때문에 신경망 학습이 제대로 이루어지지 못한다. 따라서 이 연구에서는 동일한 유체 생산량의 후보군들은 합집합하고, 각 생산량들의 합집합 결과끼리 교집합하여 각각의 생산량에 대하여 좋은 형질을 가지면서도 일정 수 이상의 후보군을 선정하였다.

채널저류층에서는 각 생산정별 오일생산량을 통하여 선정한 후보군들을 합집합하고, 마찬가지로 각 생산정별 물 생산량을 통하여

선정한 후보군들을 합집합하였다. 이 두 합집합끼리 교집합하여 다음 세대의 학습자료로 사용할 거리기반 후보군을 선정하였다(Figure 3.6(a)). PUNQ-S3 필드에서도 마찬가지로 각 생산정별 오일생산량, 정저압력, 가스-오일비를 기준으로 선정한 후보군들을 합집합하고, 이 합집합들끼리 교집합하여 다음 세대의 학습자료로 사용할 거리기반 후보군을 선정하였다(Figure 3.6(b)).



(a)



(b)

Figure 3.6 Distance-based candidate selection for proposed method:

(a) channelized reservoir, (b) PUNQ-S3 field.

4 개발기법의 검증 및 비교분석

이 연구에서 개발한 거리기반 후보군 선정과 인공신경망을 결합한 다목적 히스토리매칭 기법을 검증하기 위하여 저류층 물성분포가 비교적 단순한 셰일저류층과 저류층 물성이 정규분포를 따르지 않는 채널저류층에서 히스토리매칭을 수행하였다. 현장 적용성을 검증하기 위하여 목적함수가 많은 불균질 저류층인 PUNQ-S3 필드에서 히스토리매칭을 수행하였다.

이 장에서는 제안한 기법을 다양한 형태의 저류층에 적용하여 히스토리매칭을 수행한 결과를 소개한다. 4.1절에서는 셰일저류층과 채널저류층에서 제안한 기법의 방법론을 검증하였다. 4.2절에서는 PUNQ-S3 필드에 대한 히스토리매칭을 수행함으로써 제안한 기법의 현장 적용성을 평가하였다. 각 히스토리매칭 결과를 기존 히스토리매칭 기법인 NSGA-II, 앙상블칼만필터와 비교분석하였다.

4.1 다목적 히스토리매칭 성능분석

4.1.1 셰일저류층 히스토리매칭

이 연구에서 제안한 기법을 검증하기 위하여 한 개의 유정에서 가스만 생산되며, 지질학적 복잡성이 거의 없는 균질한 셰일저류층을 대상으로 히스토리매칭을 수행하였다. 히스토리매칭을 수행한 셰일저류층은 CMG社에서 제공한 가상 저류층이다. 북미지역에 존재하는 실제 셰일저류층의 자료를 가공하여 만든 가상 저류층으로써 4개의 수압파쇄대가 존재하는 균질 저류층이다. Figure 4.1은 해당 저류층의 모식도를 나타낸다.

셰일저류층의 생산량을 산정하기 위해 널리 쓰이는 방법에는 생산량 천이분석법(rate transient analysis)과 저류층 시뮬레이션(reservoir simulation)이 있다(kim et al., 2016). 생산량 천이분석법은 입력자료가 간단하고 미래 생산량의 신속한 추정이 가능해 유용한 방법이지만, 불균질한 저류층을 고려하기 어렵기 때문에 예측 결과의 불확실성이 크다(Anderson et al., 2012). 저류층 시뮬레이션은 복잡한 균열대 및 저류층의 불균질성 등에 대한 모사가 가능하여 정확한 미래 생산거동 분석을 위해 가장 선호되는 방법이지만, 분석을 위해 다양한 입력자료가 필요하고 연산시간이 오래 걸리는 한계가 있다(Morales-German et al., 2012). 반면, FMM (Fast Marching Method)을 이용한 유정자극범위 및 궁극가채량 예측기법은 저류층의 불균질성을 고려하면서도 빠른 시간 내에 신뢰성 있는 결과를 도출할 수 있다(Xie et al., 2015). 이 연구에서는 셰일저류층의 생산량을 산정하는 방법으로 FMM을 사용하였다.

히스토리매칭에 활용 가능한 동적자료는 Figure 4.2와 같이 5년간의 가스 생산량이다. 생산 초기에는 가스 생산량이 약 750

Mscf/day 수준이었으나 5년 후 150 Mscf/day 정도로 감소하였다. Table 4.1은 셰일저류층의 물성치를 요약한 표이다. 모암의 공극률, 유체투과도, 수압파쇄대의 길이 등을 포함한 7가지 물성이 불확실한 변수로써 히스토리매칭에 사용되었다. 총 5년간의 생산자료 중에서 2년간의 생산자료만을 이용하여 히스토리매칭을 한 후, 매칭을 완료한 2년부터 5년까지의 생산추이를 대상 셰일저류층의 생산이력과 비교하였다. 총 20세대의 히스토리매칭을 수행하였으며, 제안한 기법의 성능을 평가하기 위하여 단목적 유전 알고리즘, 앙상블칼만필터로 히스토리매칭을 수행한 결과와 비교하였다.

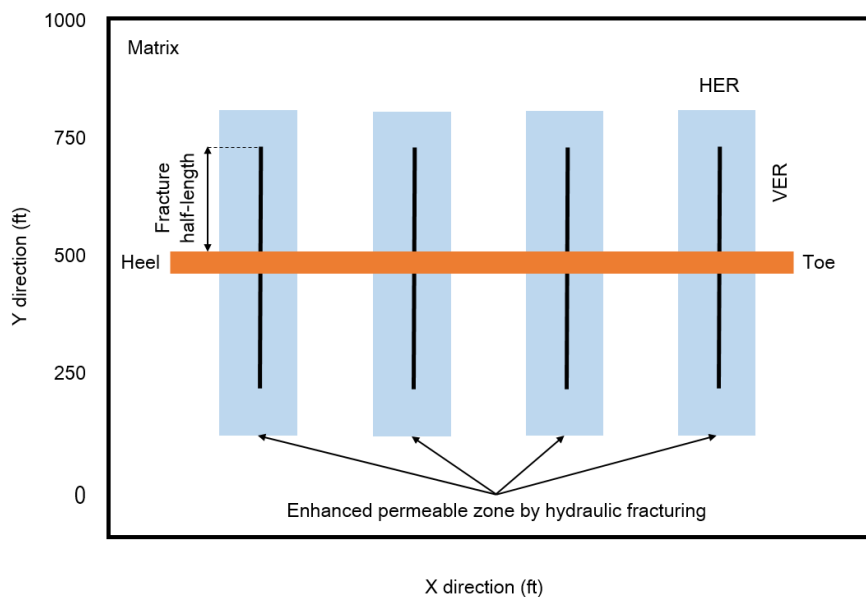


Figure 4.1 Schematic of shale reservoir with 4 hydraulic fractures.

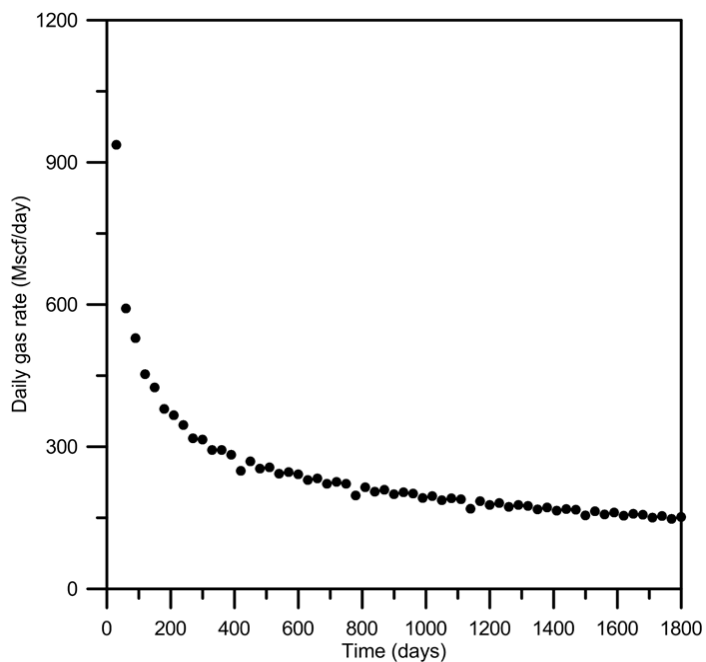


Figure 4.2 5-years gas production of the shale reservoir.

Table 4.1 Reservoir properties of the shale reservoir

Reservoir parameters (unit)	Known value	Uncertain value (minimum, maximum)
Reservoir size (x, y, z) (ft)	(1000, 1250, 450)	-
Gas viscosity (cp)	0.02	-
Gas saturation (fraction)	0.6	-
Gas-formation-volume factor (rcf/scf)	0.008	-
Initial reservoir pressure (psia)	1,500	-
Well flowing pressure (psia)	500	-
Reservoir temperature (°F)	100	-
Total compressibility (1/psia)	0.0002	-
Porosity (fraction)	-	(0.04, 0.09)
Matrix permeability (md)	-	(0.001, 0.004)
Enhanced permeability in the stimulated reservoir volume (md)	-	(0.007, 0.017)
Fracture permeability (md)	-	(0.07, 0.17)
Fracture half-length (ft)	-	(210, 270)
Horizontal Enhanced Ratio†	-	(0.05, 0.4)
Vertical Enhanced Ratio‡	-	(0.8, 1.1)

† Horizontal Enhanced Ratio= (horizontal length of enhanced zone)/(fracture length)

‡ Vertical Enhanced Ratio= (vertical length of enhanced zone)/(fracture length)

(1) 학습성능 개선 검증

히스토리매칭을 수행하기에 앞서 Table 4.1의 물성값의 범위를 이용하여 무작위로 구성된 100개의 초기모델을 생성하였다. Figure 4.3은 초기모델 100개의 생산량과 누적생산량 그래프를 나타낸다. 초기모델과 대상 저류층의 생산이력의 2년동안의 평균 오차는 일생산량 기준 29.2%, 5년 후 총생산량의 경우 26.7% 수준이다.

초기모델의 생산량과 대상 저류층의 생산이력의 차이를 정규화한 입력자료와 7개의 저류층 물성을 출력자료로 하는 인공신경망 학습을 수행하였다. 인공신경망 학습에는 MathWoRK社의 MATLAB 소프트웨어에 내장된 Neural Network Toolbox를 사용하였다. Figure 4.4는 세일저류층에서 수행한 인공신경망 학습의 모식도를 나타낸다. 2년동안의 월별 가스 생산량자료를 입력자료로 사용하였기 때문에 입력자료의 총 개수는 24개이며, 출력자료는 저류층 물성 7개이다. Table 4.2는 세일저류층 히스토리매칭을 위하여 구축한 제안한 기법 신경망의 입출력 변수를 나타낸다.

Figure 4.5(a)는 신경망학습을 통하여 생성한 첫번째 세대의 100개의 새로운 모델이다. Figure 4.3(a)와 비교하여서 실제 생산량과 가까운 모델들이 새롭게 생성되었음을 알 수 있다. 첫번째 세대의 2년동안의 오차는 일생산량 기준 13.8%이다. Figure 4.5(b)는 첫번째 세대의 모델들의 생산량과 실제 생산이력에 대하여 거리기반 군집화를 수행한 결과이다. 까만색 사각형이 생산이력의 위치이며, 생산이력과 같은 군집에 속한 43개의 모델들이(그림에서 빨간색 동그라미) 다음세대의 학습을 위한 거리기반 후보군으로 선정되었다. Figure 4.5(c)는 선정된 43개의 모델들을 도시한 그래프이다. Figure 4.5(a)와 비교하여서 실제 생산이력과 가까운 모델들이 선정되었음을

알 수 있다. Figure 4.5(d)는 Figure 4.5(c)에서 선정된 43개의 모델들을 새로운 학습자료로 하는 신경망학습을 통하여 생성한 두번째 세대의 100개의 모델이다. Figure 4.5(a)와 비교하여 실제 생산량과 가까운 모델들이 생성되었음을 알 수 있다.

20세대까지 히스토리매칭을 수행하였으며, Figure 4.6과 Figure 4.7은 히스토리매칭을 수행한 각 세대의 생산량과 누적생산량을 각각 도시한 그래프이다. 세대가 지나갈수록 실제 생산량에 가까운 결과를 보여주었다. Table 4.3은 Figure 4.6과 Figure 4.7에 도시된 각 세대의 히스토리매칭 구간 동안의 평균 생산량 오차와 예측 구간의 생산량 오차, 마지막 예측일의 평균 누적생산량 오차를 요약한 표이다. 세대가 지날수록 세가지 오차가 모두 감소하였으며, 이와 같은 결과를 통하여 제안한 기법을 통해 신경망의 학습성능이 개선됨을 확인할 수 있다.

Table 4.2 Neuron description of neural network for shale reservoir

Description of the neuron	Number of neurons and algorithm
Input neuron (gas production rate)	24
Hidden neuron [Multi-layer]	[500, 500]
Output neuron (reservoir properties)	7
Train algorithm	Trainidx†

† Gradient descent with momentum and adaptive learning rate backpropagation

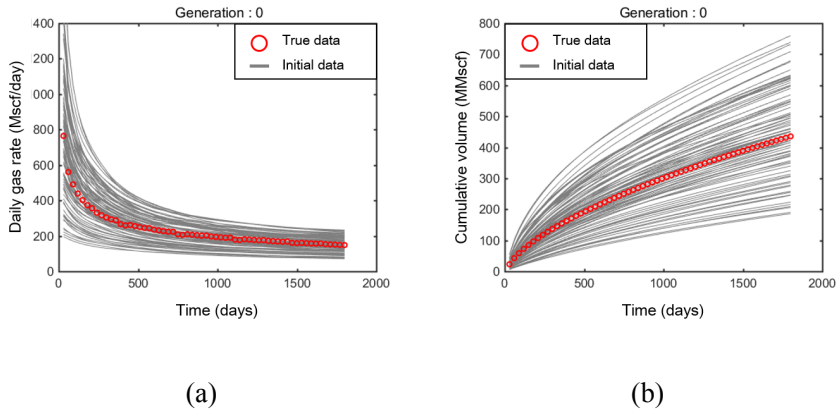


Figure 4.3 Production profile of initial models:
(a) daily gas rate, (b) cumulative volume.

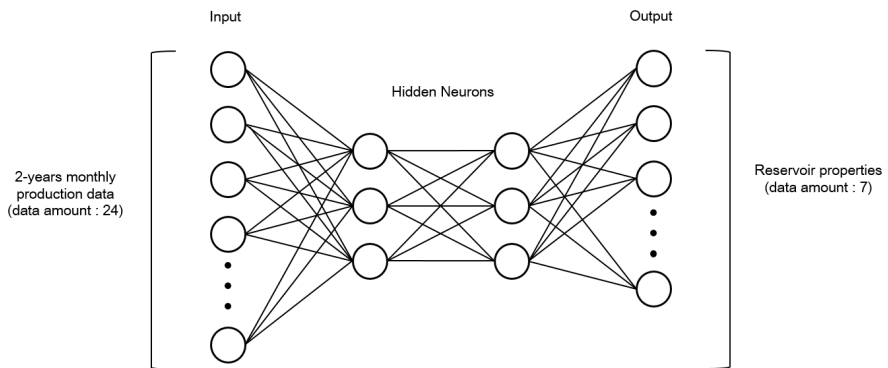
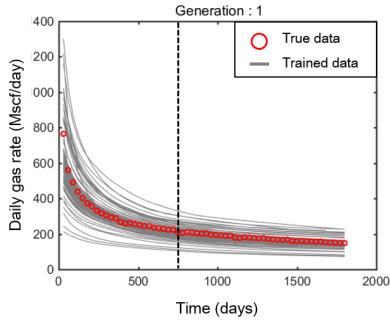
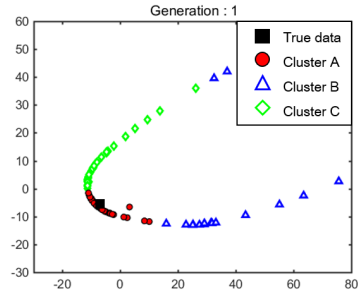


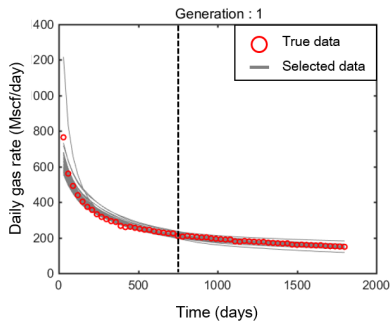
Figure 4.4 ANN training for the shale reservoir.



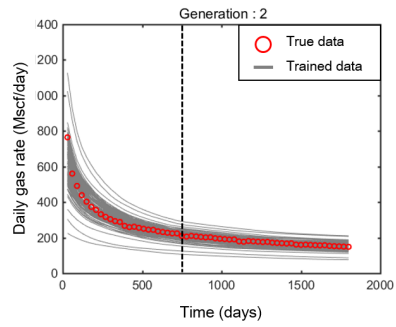
(a)



(b)



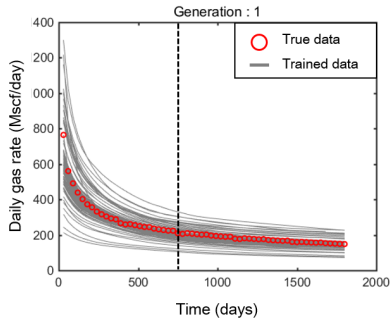
(c)



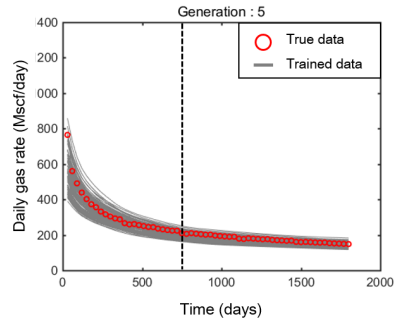
(d)

Figure 4.5 Production profile of trained models:

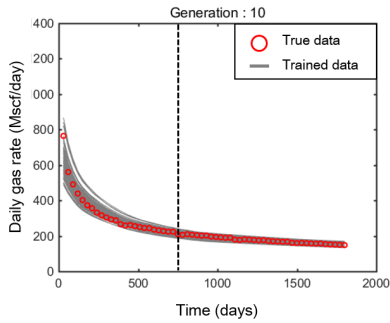
- (a) models trained using initial models,
- (b) the selection of cluster involving the true dataset,
- (c) selected models, (d) models trained using selected models.



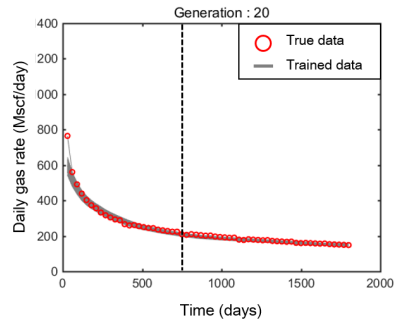
(a)



(b)

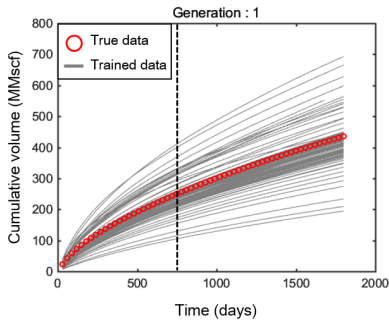


(c)

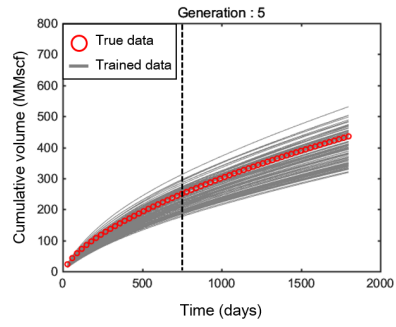


(d)

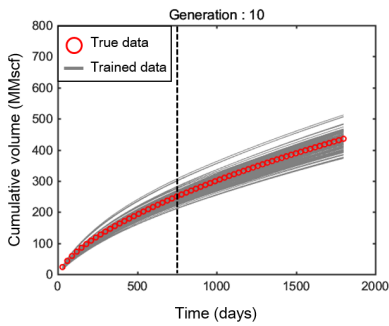
Figure 4.6 Production profiles of trained models: (a) 1st generation, (b) 5th generation, (c) 10th generation, (d) 20th generation.



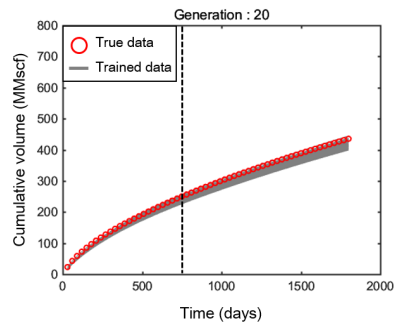
(a)



(b)



(c)



(d)

Figure 4.7 Cumulative volume of trained models: (a) 1st generation, (b) 5th generation, (c) 10th generation, (d) 20th generation.

Table 4.3 Errors of initial and trained models

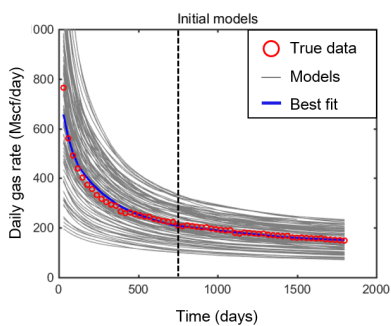
Error	Initial	Generation: 1	Generation: 5	Generation: 10	Generation: 20
Daily gas rate (matching period)	29.2%	13.8%	11.5%	7.1%	4.3%
Daily gas rate (prediction period)	25.3%	7.5%	6.2%	2.9%	1.8%
Cumulative volume (1800 days)	26.7%	13.4%	11.3%	5.5%	3.5%

(2) 히스토리매칭 결과 비교분석

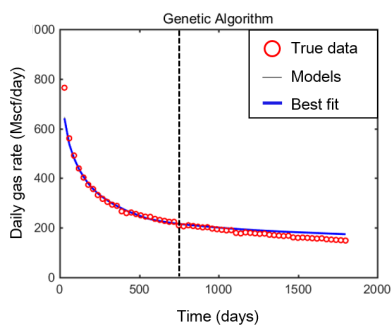
제안한 기법을 통해 히스토리매칭을 수행한 결과와 기존 기법으로 히스토리매칭을 수행한 결과를 비교하였다. 비교에 사용한 기존 기법은 단목적 유전 알고리즘과 앙상블기반기법이다. 제안한 기법의 최종결과는 마지막 세대에서 선정된 거리기반 후보군을 사용하였다. 각 기법은 동일하게 Table 4.1에 있는 7개의 불확실한 저류층 물성에 대하여 히스토리매칭을 수행하였으며, Figure 4.2의 초기모델을 사용하였다. 유전 알고리즘의 경우 제안한 기법과 동일하게 20세대까지 히스토리매칭을 수행하였고, 앙상블칼만필터는 동적자료로 2년간의 월별 가스생산량을 이용하여 24번 교정하였다.

Figure 4.8과 Figure 4.9는 초기 모델과 각 기법으로 히스토리매칭을 수행한 최종 결과를 비교한 일별가스생산량, 누적생산량 그래프이다. 빨간색 동그라미가 실제 생산이력을 의미하며, 회색 선은 각각의 모델들의 생산량, 파란색 선은 마지막 예측일 누적생산량 기준으로 가장 오차가 작은 모델의 생산량이다. 세로로 된 점선은 히스토리매칭 구간인 720일을 의미한다. Table 4.4는 각 기법의 예측결과를 평균 오차를 통해서 비교한 표이다. 유전 알고리즘과 앙상블칼만필터, 제안한 기법 모두 초기 모델이 비하여 개선된 예측성능을 보여주었다. Figure 4.10은 각 기법의 마지막 예측일 누적생산량과 실제 누적생산량의 비를 박스플롯으로 나타낸 그림이다. 유전 알고리즘은 예측구간에서 오차는 작았지만, 실제 생산이력과는 벗어난 결과를 보여주었다. 이는 단목적 유전 알고리즘의 특징 때문인 것으로 분석할 수 있다. 단목적 유전 알고리즘은 오차가 가장 작은 해를 최적해로 탐색해나가는 과정에서 단일해로 빠르게 수렴하는 경향을 보인다. 이로 인해 해의 다양성을 보존하지 못하여

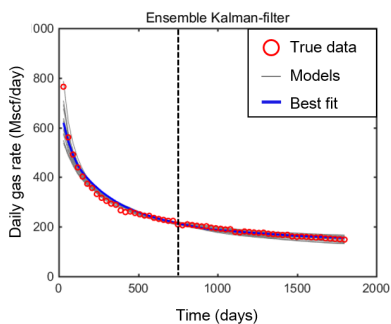
불확실성 평가를 수행하지 못할 뿐만 아니라, 국소해(local minimum)에 빠질 위험이 있다. 앙상블칼만필터는 예측범위 안에 실제 생산이력을 포함하기는 하지만, 제안한 기법보다는 큰 오차를 보였다. 제안한 기법이 전체적으로 가장 좋은 예측성능을 보였지만, 저류층 물성분포가 비교적 단순하며 목적함수가 1개인 셰일저류층에서는 다른 히스토리매칭 기법들도 초기모델에 비하여 개선된 히스토리매칭 결과를 보여주었다.



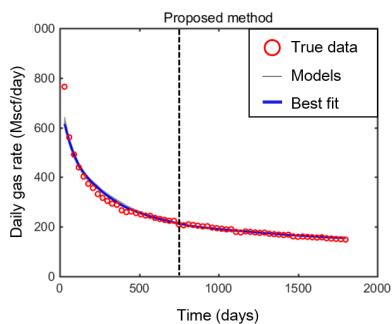
(a)



(b)

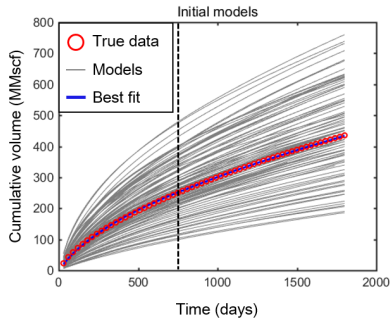


(c)

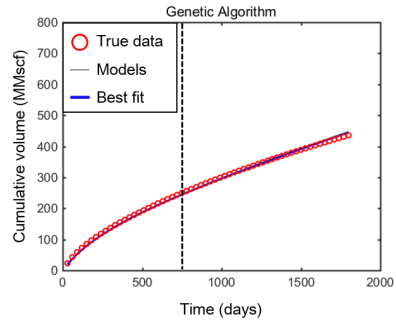


(d)

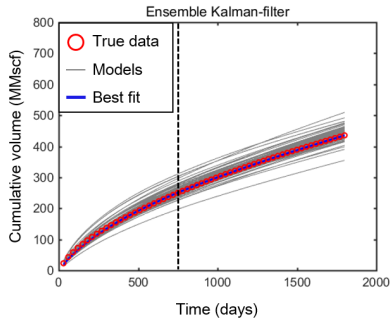
Figure 4.8 Comparison of history matching results for shale reservoir (production profile): (a) initial models, (b) Genetic Algorithm, (c) Ensemble Kalman-filter, (d) proposed method.



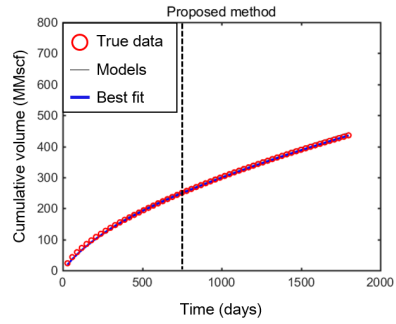
(a)



(b)



(c)



(d)

Figure 4.9 Comparison of history matching results for shale reservoir (cumulative volume): (a) initial models, (b) Genetic Algorithm, (c) Ensemble Kalman-filter, (d) proposed method.

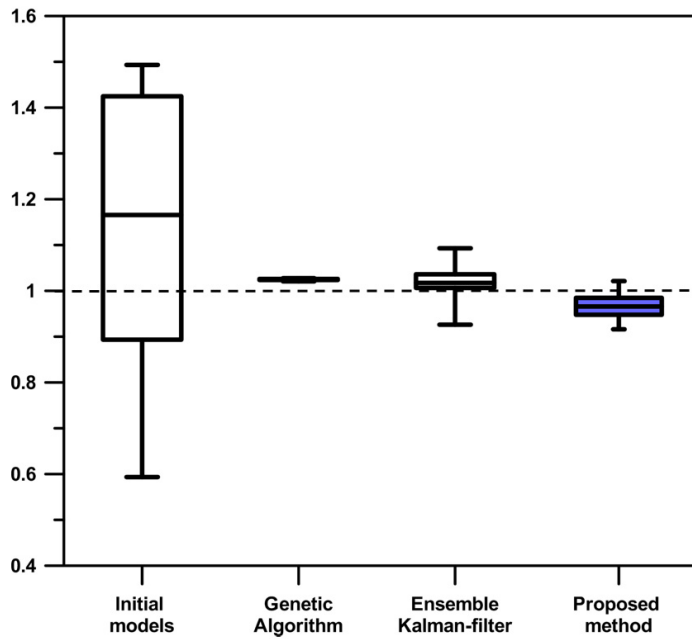


Figure 4.10 Box plot of history matching results.

Table 4.4 Errors of history matching results

Error	Initial models	Genetic Algorithm	Ensemble Kalman-filter	Proposed method
Daily gas rate (matching period)	29.2%	2.5%	6.9%	4.3%
Daily gas rate (prediction period)	25.3%	6.5%	5.4%	1.8%
Cumulative volume (1800 days)	26.7%	2.5%	3.7%	3.5%

4.1.2 채널저류층 다목적 히스토리매칭

저류층 물성분포가 정규분포를 따르지 않는 불균질 저류층에서의 히스토리매칭 결과를 확인하기 위하여 채널저류층을 대상으로 히스토리매칭을 수행하였다. 스탠포드 대학교의 SGeMS 소프트웨어를 이용하여 채널저류층의 참조모델과 초기 모델들을 생성하였다.

먼저 SGeMS의 TIgenerator 모듈을 사용하여 Figure 4.11(a)와 같이 채널의 방향이 0도인 트레이닝 이미지(training image)를 생성하였다. 이 트레이닝 이미지와 SGeMS의 SNESim 시뮬레이션을 이용하여 Figure 4.11(b)와 같이 참조모델을 생성하였다. Table 4.5는 트레이닝 이미지를 생성하기 위해 설정한 채널의 폭, 길이, 방향성 등에 관한 인자를 나타내며, Table 4.6은 트레이닝 이미지로부터 참조모델과 초기모델을 생성하기 위해 사용한 인자들을 나타낸다.

Case 1은 참조모델과 같은 트레이닝 이미지를 사용하여 100개의 초기모델을 생성하였다. Case 2는 참조모델과 다른 트레이닝 이미지로 초기모델을 생성하였다. 즉, 채널의 방향이 45도인 TI에서 34개, 방향이 90도인 TI에서 33개, 방향이 135도인 TI에서 33개, 총 100개의 초기모델을 생성하였다(Figure 4.12). Figure 4.13은 참조모델의 유체투과도 분포도와 유정의 위치를 나타낸다. Table 4.7과 같이 4개의 생산정과 1개의 주입정이 있으며, 저류층 특성은 Table 4.8과 같다. 4개의 생산정에서의 월별 오일, 물 생산량, 총 8개의 목적함수에 대하여 히스토리매칭을 수행하였다.

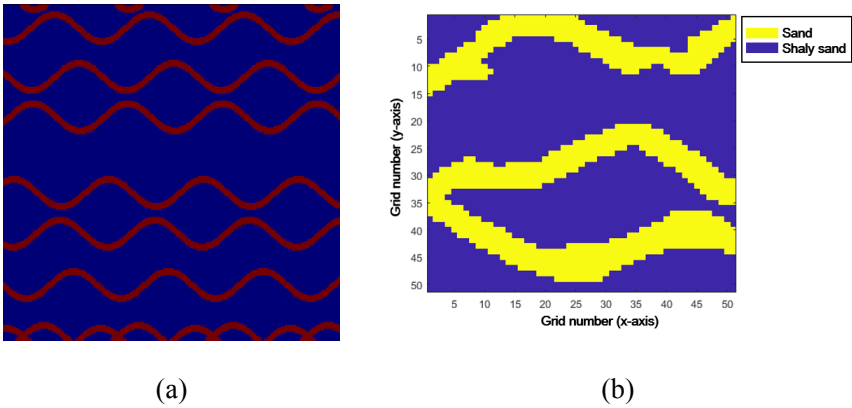


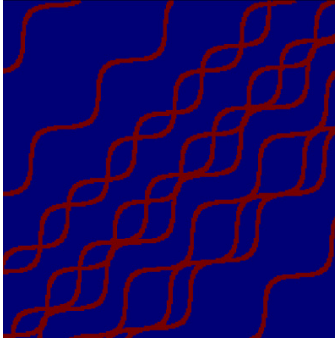
Figure 4.11 Reference field: (a) training image, (b) field.

Table 4.5 Parameters for training image

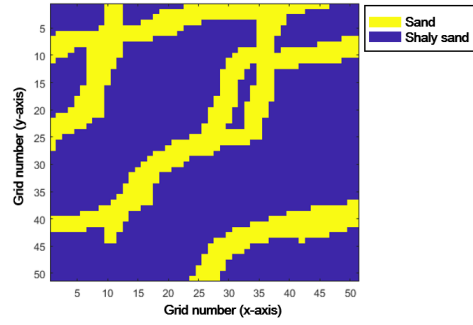
Parameters	Value
Grid system (number of x, y, z)	250 x 250 x 1
Geobody type	Sinusoid
Length	10,000
Width	5
Orientation	(0, 45, 90, 135)
Amplitude	10
Wave Length	70
Sand ratio	0.25

Table 4.6 Parameters for SNESim simulation

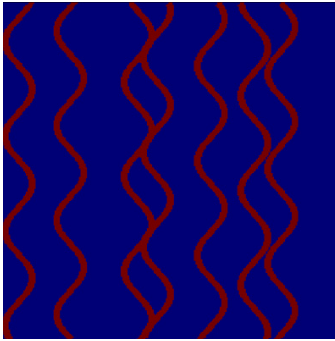
Parameters	Values
Grid system (number of x, y, z)	51 x 51 x 1
Target marginal distribution	Sand=0.6, Shaly sand=0.4
Search template (Nodes)	60
Search template (Ranges)	30



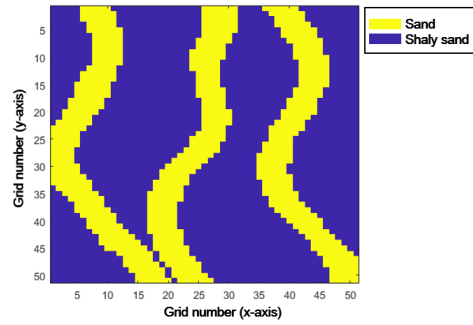
(a)



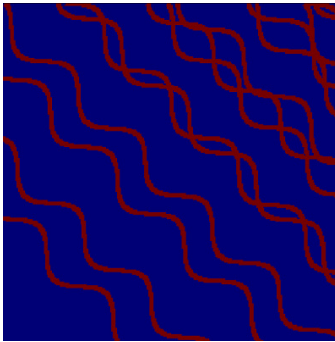
(b)



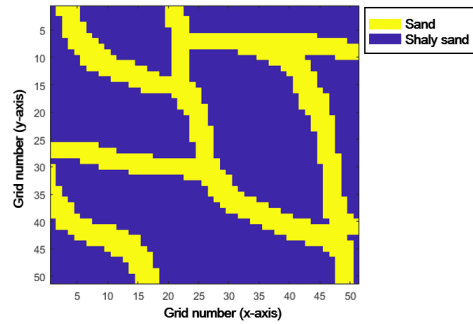
(c)



(d)



(e)



(f)

Figure 4.12 Initial models for case 2:

(a) training image (orientation: 45), (b) field (orientation: 45),
(c) training image (orientation: 90), (d) field (orientation: 90),
(e) training image (orientation: 135), (f) field (orientation: 135).

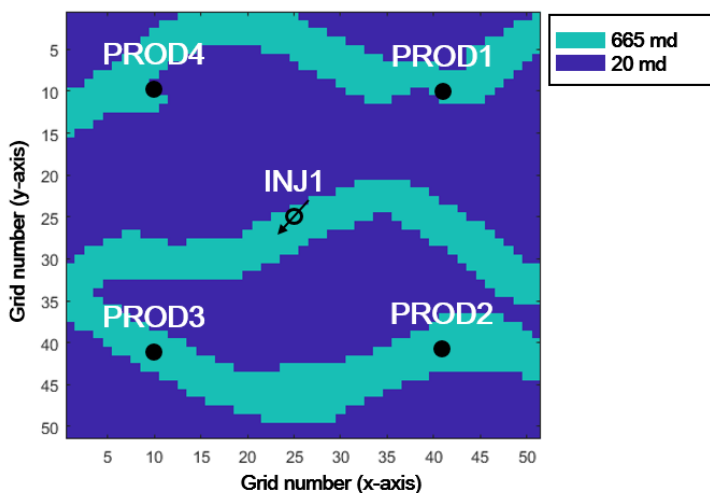


Figure 4.13 Permeability distribution of reference field.

Table 4.7 Well information and boundary conditions

Well name	Location (X, Y)	Control mode	Facies
PROD1	(41, 10)	300 psia	Sand
PROD2	(41, 41)	300 psia	Sand
PROD3	(10, 41)	300 psia	Sand
PROD4	(10, 10)	300 psia	Sand
INJ1	(25, 25)	400 STB/day	Sand

Table 4.8 Reservoir properties of channelized reservoir.

Reservoir parameters (unit)	Values
Grid system(number of x, y, z)	51 x 51 x 1
Width, ft	1,500
Length, ft	1,500
Thickness, ft	40
Porosity, fraction	0.3
Initial pressure, psia	2,000
Permeability of sand, md	665
Permeability of shaly sand, md	20
Formation volume factor of water, rb/STB	1 at 2,000 psig
Formation volume factor of oil, rb/STB	1.01 at 2,000 psig
Density of water, lb/ ft ³	48.6
Density of oil, lb/ ft ³	62.3
Viscosity of water, cp	1
Viscosity of oil, cp	3
Compressibility of water, psi ⁻¹	5.0E-07 at 2,000 psig
Compressibility of oil, psi ⁻¹	3.0E-05 at 2,000 psig

(1) 심층신경망 은닉층 선정

제안한 기법으로 히스토리매칭을 수행하기에 앞서 $50 \times 50 \times 1 = 2,601$ 개 격자의 유체투과도를 100개의 출력층 뉴런으로 압축하는 심층 오토인코더를 구축하였다. 심층 오토인코더 학습에는 MathWorK社의 MATLAB 소프트웨어에 내장된 Neural Network Toolbox를 사용하였다. 심층신경망의 안정성을 확보하기 위하여 두 번의 부호화 과정을 거쳤으며, 첫 번째 부호화 은닉층의 뉴런수에 대한 민감도 분석을 통하여 최적의 뉴런수를 선정하였다.

Table 4.9는 첫번째 부호화 은닉층 뉴런수를 각각 200개, 300개, 500개, 1,000개로 설정하였을 때의 학습오차와 연산시간을 비교한 표이다. 뉴런의 수가 많을수록 학습오차는 줄어들지만, 연산시간이 증가하는 것을 알 수 있다. Figure 4.14는 첫 번째 부호화 은닉층의 뉴런수에 따른 총 학습시간과 총 학습오차를 나타낸 그래프이다. 200개에서 300개로 뉴런수가 증가하면 학습오차는 크게 줄어들었지만, 학습시간은 크게 달라지지 않았다. 반면에 300개에서 1,000개로 뉴런수가 증가하면 학습오차는 크게 달라지지 않지만, 학습시간은 4배정도 차이가 났다. 이 연구에서는 첫 번째 부호화 은닉층의 최적 뉴런수를 300개로 선정하였고, Table 4.10은 채널저류층 히스토리매칭을 위하여 구축한 심층신경망의 입출력 변수를 나타낸다. Table 4.11은 심층신경망 기법을 통해 부호화된 출력층을 이용하여 구축한 채널저류층의 신경망 입출력 변수를 나타낸다.

Table 4.9 Training error and training time depending on the number of hidden neurons for channelized reservoir

Number of hidden neurons for first encoding	200	300	500	1,000
Training error for first encoding	2.0%	0.7%	0.5%	0.3%
Training time for first encoding	364.2 sec	541.6 sec	859.1 sec	1859.8 sec
Training error for second encoding	2.1%	0.8%	0.5%	0.3%
Training time for second encoding	118.6 sec	115.7 sec	153.0 sec	575.9 sec



Figure 4.14 Training error and training time depending on the number of hidden neurons for channelized reservoir.

Table 4.10 Neuron description of autoencoder for channelized reservoir

Description of the neuron	Number of neurons and algorithm
Input neuron (grid permeability)	$51 \times 51 = 2601$
Hidden neuron	[300, 100]
Train algorithm	Trainscg†

† Scaled conjugate gradient backpropagation

Table 4.11 Neuron description of neural network for channelized reservoir

Description of the neuron	Number of neurons and algorithm
Input neuron (oil rate, water rate)	$24 \times 4 \times 2 = 192$
Hidden neuron [Multi-layer]	[500, 500]
Output neuron (Encoded grid permeability)	100
Train algorithm	Traingdx††

†† Gradient descent with momentum and adaptive learning rate backpropagation

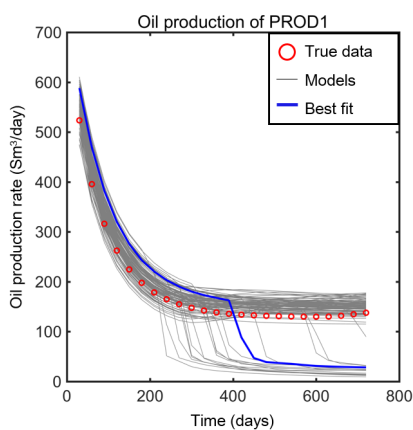
(2) Case 1: 참조모델과 같은 트레이닝 이미지를 사용한 경우

Figure 4.15는 참조모델과 트레이닝 이미지가 같은 초기모델 100개의 720일까지의 오일 생산이력을 나타낸다. Schulumberger社의 ECLIPSE 100 소프트웨어를 통하여 저류층 시뮬레이션을 수행하였다. 빨간색 동그라미는 참조모델의 생산이력, 회색 선은 초기모델의 생산이력, 파란색 선은 720일의 총 오일생산량의 오차가 가장 작은 모델의 생산량을 나타낸다. 참조모델의 생산이력을 보면, INJ1과 사암층으로 연결된 PROD3에서만 720일 이전에 물돌파가 일어나서 오일 생산량이 감소한 것을 관찰할 수 있다. 제안한 기법의 성능을 평가하기 위하여 제안한 기법을 통해 히스토리매칭을 수행한 결과와 기존 기법으로 히스토리매칭을 수행한 결과를 비교하였다. 비교에 사용한 기존 기법은 NSGA-II와 앙상블 칼만필터로써 개발한 기법과 동일한 초기모델을 사용하여 히스토리매칭을 수행하였다.

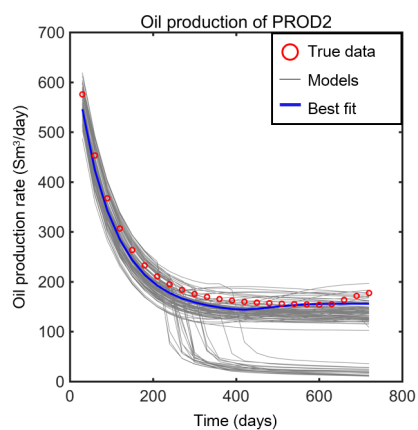
Figure 4.16은 참조모델과 초기모델, 각 히스토리매칭 기법으로 매칭된 대표모델의 유체투과도 분포도를 보여준다. 각 기법에서의 대표모델은 1,800일에서의 총 오일생산량 기준으로 오차가 가장 작은 모델을 선정하였다. 참조모델에서는 주입정과 PROD3, PROD2가 사암채널로 연결되어 있고, PROD1과 PROD4는 주입정과 연결되어 있지 않다. 초기모델의 대표모델은 주입정과 PROD1, PROD4가 연결되어 있으며, PROD3, PROD2와는 단절되어 있는 모습을 보인다. NSGA-II의 대표모델은 전체적인 경향성은 유사하지만, 주입정과 PROD3이 단절되어 있으며, 앙상블 칼만필터의 대표모델은 초기모델에 비해서 크게 교정되지 못한 모습을 보인다. 반면에, 제안한 기법의 대표모델은 INJ1과 PROD3의 연결성을 보여주면서도 전체적인 채널 모양을 비교적 잘 보여주고 있다.

Figure 4.17은 각 기법으로 매칭된 최종 모델들과 참조모델의 평균 유체투과도 오차를 나타낸다. 하얀색에 가까울수록 오차가 낮아서 매칭이 잘된 부분이며, 까만색에 가까울수록 오차가 커서 매칭이 잘 안된 부분이다. 초기모델은 각 유정이 위치한 부분을 제외하고는 전체적으로 큰 오차를 보이고 있음을 알 수 있다. NSGA-II와 앙상블 칼만필터도 초기모델에 비하여 크게 개선되지 못하는 모습을 보이고 있다. 반면에, 제안한 기법은 다른 기법에 비하여 밝은 부분이 훨씬 많이 보이며 대체적으로 매칭이 잘 이루어졌음을 확인할 수 있었다. Figure 4.18-4.19는 각 기법으로 매칭된 최종 모델들의 PROD3, PROD2의 오일과 물 생산량 그래프를 보여준다. 빨간색 동그라미는 참조모델의 생산이력, 회색선은 최종 모델들의 생산량, 파란색 선은 대표모델의 생산량을 의미한다. 세로로 된 검정색 점선은 매칭기간인 720일을 보여준다.

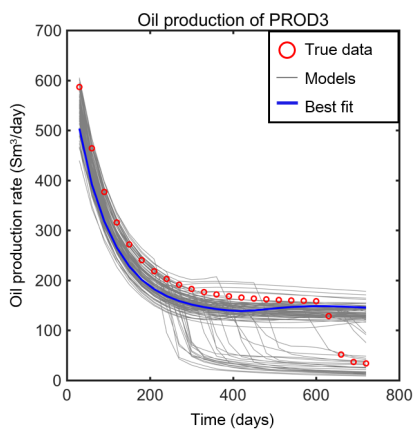
앙상블칼만필터의 최종 모델은 생산정에서의 물돌파(water breakthrough) 시점이 참조모델과 크게 다른 것을 볼 수 있다. 이는 앙상블 칼만필터의 기본가정인 정규분포를 따르지 않는 물성분포의 저류층에서는 제대로 교정이 이루어지지 못하는 앙상블 칼만필터의 한계점을 보여준다. NSGA-II의 최종 모델은 매칭기간 내에 일어나는 PROD3의 물돌파 시점은 예측하였지만, 예측기간에 일어나는 PROD2의 물돌파 시점은 제대로 예측하지 못하였다. 반면에, 제안한 기법의 최종모델은 매칭기간 내에 일어나는 PROD3의 물돌파 시점과 예측기간에 일어나는 PROD2의 물돌파 시점 모두 참조모델과 유사하게 예측하였다.



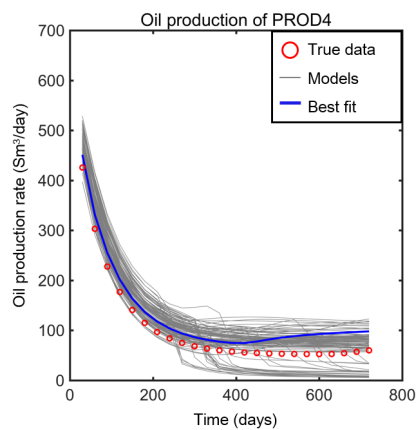
(a)



(b)

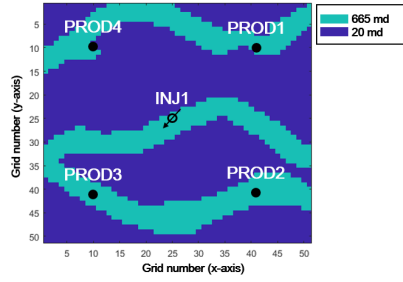


(c)

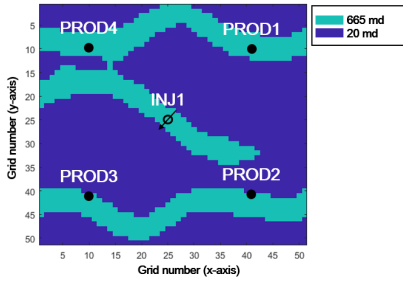


(d)

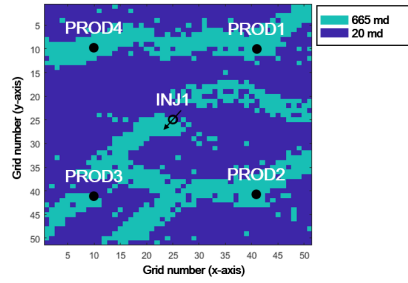
Figure 4.15 Oil production profiles of initial models until 720 days (case 1): (a) PROD1, (b) PROD2, (c) PROD3, (d) PROD4.



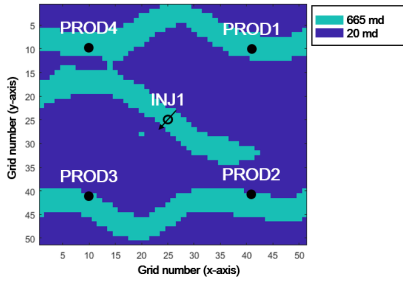
(a)



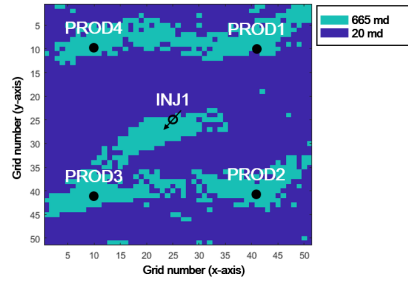
(b)



(c)



(d)

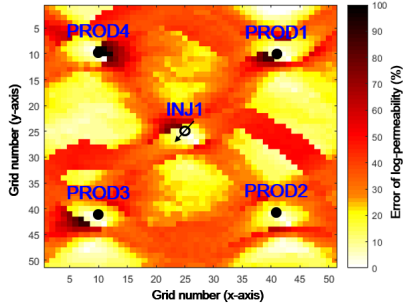


(e)

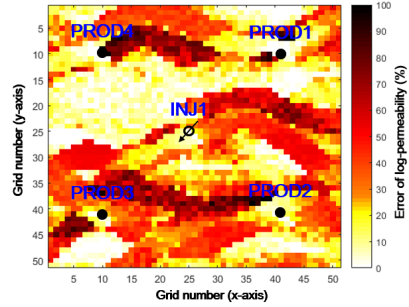
Figure 4.16 Permeability distribution of best fit model (case 1):

(a) reference field, (b) initial models,

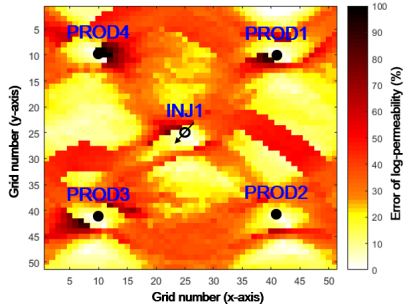
(c) NSGA-II, (d) Ensemble Kalman-filter, (e) proposed method.



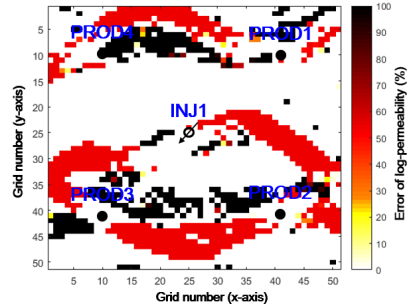
(a)



(b)



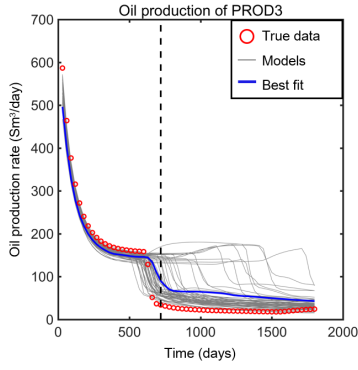
(c)



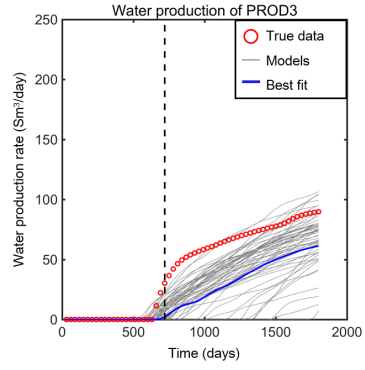
(d)

Figure 4.17 Average error map of permeability (case 1):

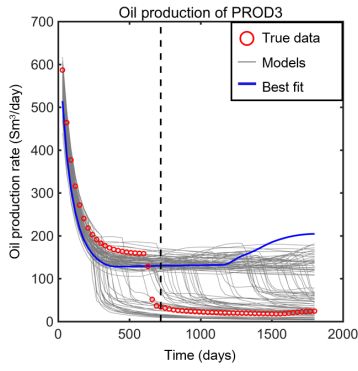
(a) initial models, (b) matched models by NSGA-II, (c) assimilated models by Ensemble Kalman-filter, (d) selected models by proposed method.



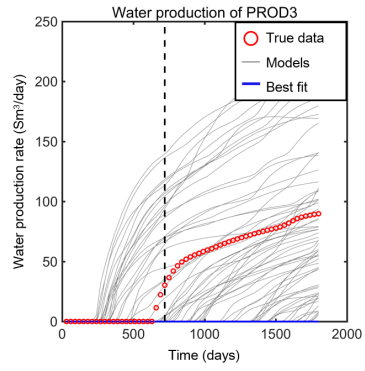
(a)



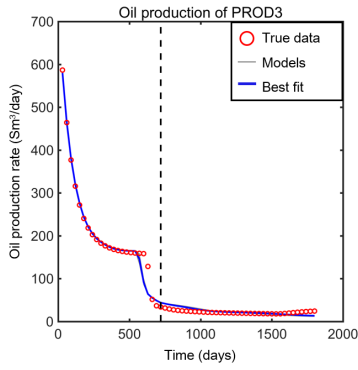
(b)



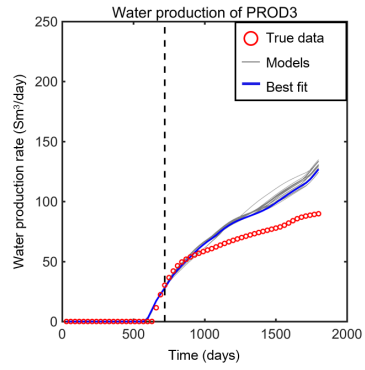
(c)



(d)



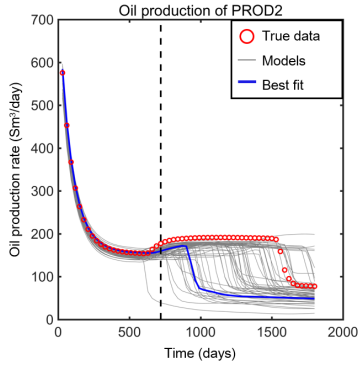
(e)



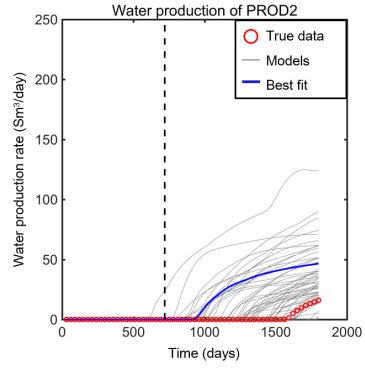
(f)

Figure 4.18 Production profiles of PROD3 (case 1):

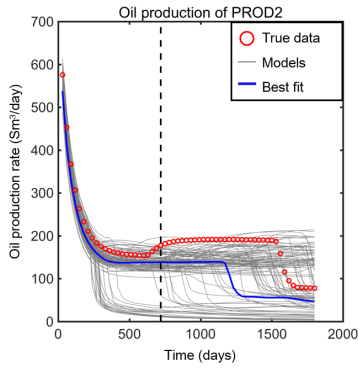
(a) (b) NSGA-II (oil, water production), (c) (d) Ensemble Kalman-filter (oil, water production), (e) (f) proposed method (oil, water production).



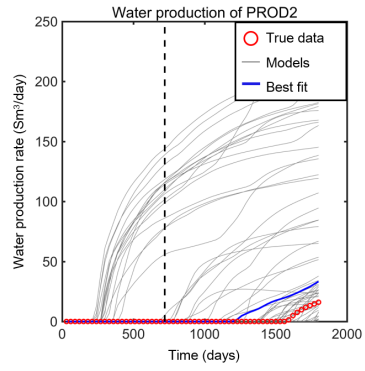
(a)



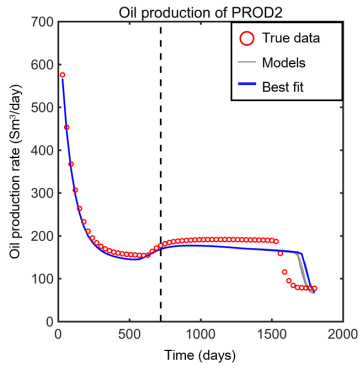
(b)



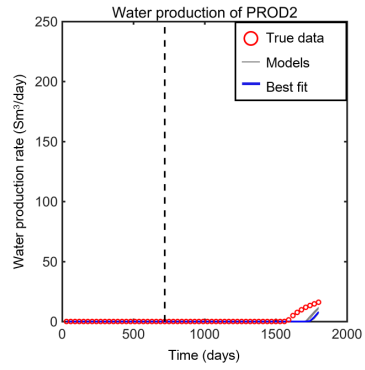
(c)



(d)



(e)



(f)

Figure 4.19 Production profiles of PROD2 (case 1):

(a) (b) NSGA-II (oil, water production), (c) (d) Ensemble Kalman-filter (oil, water production), (e) (f) proposed method (oil, water production).

Table 4.12는 각 기법을 통해 히스토리매칭을 수행한 결과들의 매칭구간 생산정별 평균 오일 생산량 오차를 요약한 표이다. 제안한 기법은 매칭구간에서의 각 생산정의 물돌과 시점을 잘 예측하였기 때문에 모든 생산정에서 가장 작은 오차를 보여주었으며, 앙상블칼만필터는 초기모델에 비하여 오차가 크게 개선되지 못한 모습을 보여주었다. Table 4.13은 각 기법을 통해 히스토리매칭을 수행한 결과들의 예측구간 생산정별 평균 오일 생산량 오차와 총 오일 생산량, 총 물 생산량 오차를 요약한 표이다. 앙상블칼만필터는 매칭구간과 마찬가지로 예측구간에서도 초기모델에 비하여 크게 개선되지 못하는 모습을 보여주었으나, NSGA-II는 초기모델이 비하여 오차가 어느 정도 감소하는 모습을 보여주었다. 하지만, PROD3의 물돌과 시점을 제대로 예측하지 못하여 PROD3에서의 생산량 오차가 매우 크게 나타났다. 제안한 기법은 모든 생산정의 물돌과 시점을 비교적 정확하게 예측함으로써 매칭구간과 마찬가지로 예측구간에서도 가장 낮은 오차를 보였다.

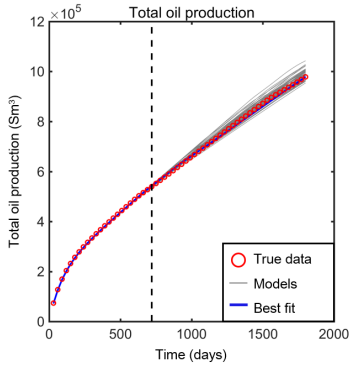
Figure 4.20은 각 기법을 통해 히스토리매칭을 수행한 결과들의 총 오일생산량과 총 물생산량을 나타낸 그래프이며, Figure 4.21은 박스플롯으로 나타낸 그림이다. 초기모델과 앙상블 칼만필터는 총 오일 생산량을 과소예측하고, 총 물 생산량을 과대예측하는 경향을 보여주면서 예측에 실패한 모습을 보여주었다. NSGA-II와 제안한 기법은 참조모델의 총 오일 생산량, 총 오일생산량을 잘 예측하였으며, 제안한 기법이 NSGA-II 보다 참조모델과 더 유사하게 생산량을 예측하였다.

Table 4.12 Errors of history matching results for case 1
(during matching period, 720 days)

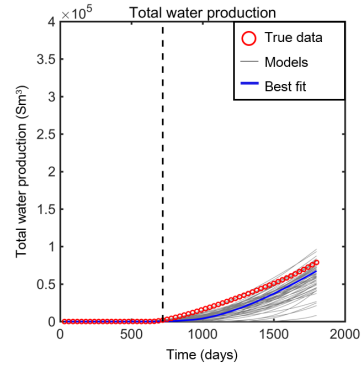
method	Initial models	NSGA-II	Ensemble Kalman-filter	Proposed method
PROD1	18.5%	5.2%	15.1%	1.1%
PROD2	21.2%	5.0%	19.0%	4.8%
PROD3	42.6%	29.7%	41.8%	8.0%
PROD4	41.8%	15.9%	40.6%	11.3%

Table 4.13 Errors of history matching results for case 1
(during prediction period, 1,800 days)

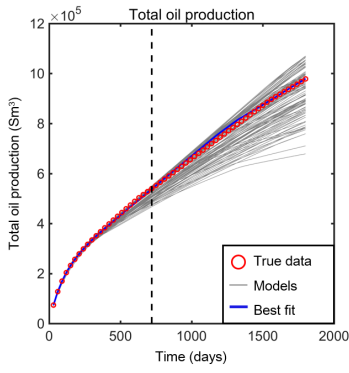
method	Initial models	NSGA-II	Ensemble Kalman-filter	Proposed method
PROD1	28.3%	10.6%	25.4%	2.2%
PROD2	37.9%	24.9%	35.4%	13.4%
PROD3	218.8%	117.6%	209.5%	15.5%
PROD4	44.9%	25.9%	43.9%	10.0%
Total oil production	10.2%	1.7%	8.5%	0.7%
Total water production	133.5%	26.5%	108.9%	20.6%



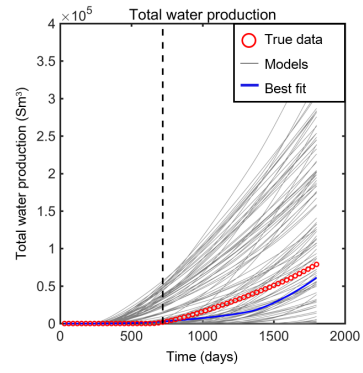
(a)



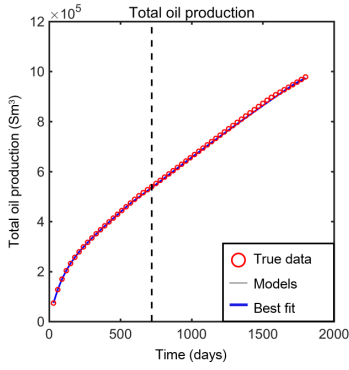
(b)



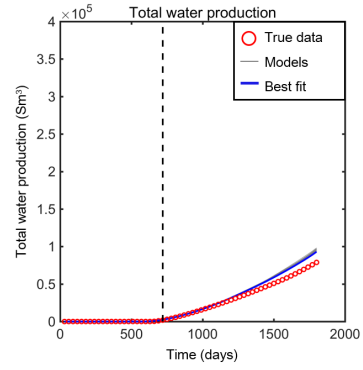
(c)



(d)



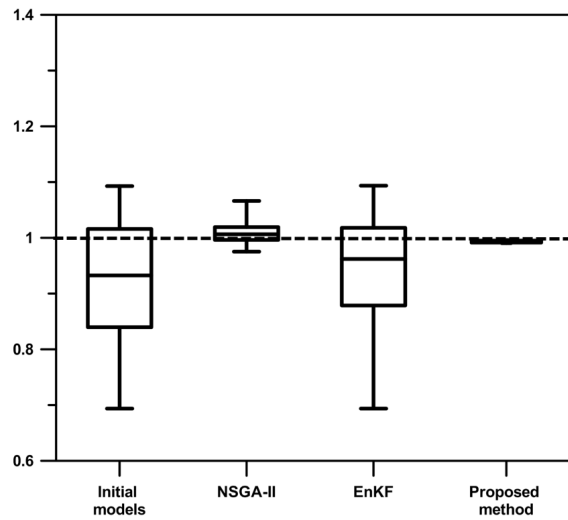
(e)



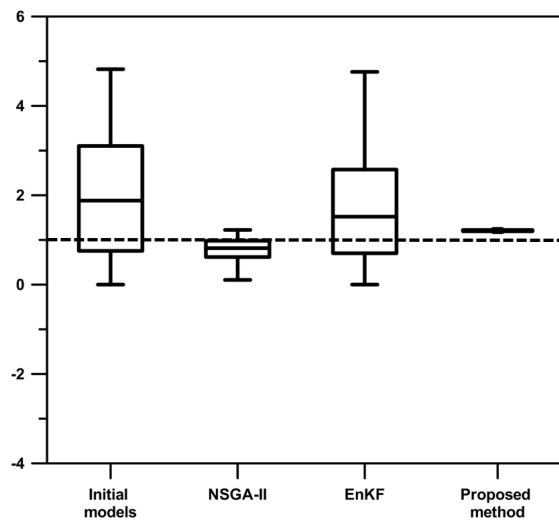
(f)

Figure 4.20 Total production (case 1):

- (a) (b) NSGA-II (total oil, water production),
(c) (d) Ensemble Kalman-filter (total oil, water production),
(e) (f) proposed method (total oil, water production).



(a)



(b)

Figure 4.21 Box plot of history matching results at 1,800 days (case 1):
(a) Total oil production, (b) Total water production.

(3) Case 2: 참조모델과 다른 트레이닝 이미지를 사용한 경우

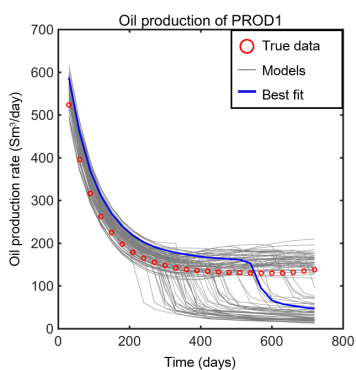
초기모델이 지질학적 불확실성이 큰 경우에도 신뢰성 있는 히스토리매칭 결과를 산출함을 확인하기 위하여 초기모델 생성시 참조모델과 다른 트레이닝 이미지를 사용하였다. Figure 4.22는 참조모델과 트레이닝 이미지가 다른 초기모델 100개의 720일까지의 오일 생산이력을 나타낸다. 빨간색 동그라미는 참조모델의 생산이력, 회색 선은 초기모델의 생산이력, 파란색 선은 720일의 총 오일생산량의 오차가 가장 작은 모델의 생산량을 나타낸다.

Figure 4.23은 참조모델과 초기모델, 각 히스토리매칭 기법으로 매칭된 대표모델의 유체투과도 분포도를 보여준다. 각 기법에서의 대표모델은 1,800일에서의 총 오일생산량 기준으로 오차가 가장 작은 모델을 선정하였다. 참조모델에서는 INJ1과 PROD3, PROD2가 사암채널로 연결되어 있고, PROD1과 PROD4는 주입정과 연결되어 있지 않다. 초기모델의 대표모델은 참조모델과 다른 트레이닝 이미지에서 생성되었기 때문에, 참조모델과는 다른 수직방향의 채널양상을 보인다. NSGA-II의 대표모델은 사암이 채널 모양으로 연결되어 있지 않고, 따로 분포되는 경향이 보이며, 앙상블 칼만필터의 대표모델은 초기모델에 비해서 크게 교정되지 못한 모습을 보인다. 반면에, 제안한 기법의 대표모델은 INJ1과 PROD3의 연결성을 보여주면서도 채널 모양을 비교적 잘 보여주고 있다.

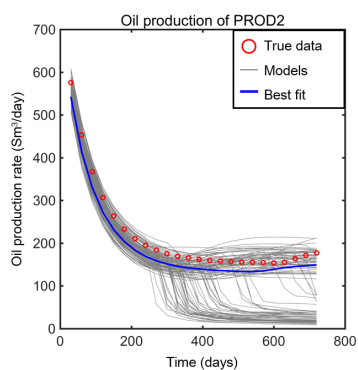
Figure 4.24은 각 기법으로 매칭된 최종 모델들과 참조모델의 평균 유체투과도 오차를 나타낸다. 하얀색에 가까울수록 오차가 낮아서 매칭이 잘된 부분이며, 까만색에 가까울수록 오차가 커서 매칭이 잘 안된 부분이다. 초기모델은 각 유정이 위치한 부분을 제외하고는 전체적으로 큰 오차를 보이고 있음을 알 수 있다. NSGA-II와 앙상블

칼만필터도 초기모델에 비하여 크게 개선되지 못하는 모습을 보이고 있다. 반면에, 제안한 기법은 다른 기법에 비하여 밝은 부분이 훨씬 많이 보이며 대체적으로 매칭이 잘 이루어졌음을 확인할 수 있었다. Figure 4.25-4.26은 각 기법으로 매칭된 최종 모델들의 PROD3, PROD2의 오일과 물 생산량 그래프를 보여준다. 빨간색 동그라미는 참조모델의 생산이력, 회색선은 최종 모델들의 생산량, 파란색 선은 대표 모델의 생산량을 의미한다. 세로로 된 검정색 점선은 매칭기간인 720일을 보여준다.

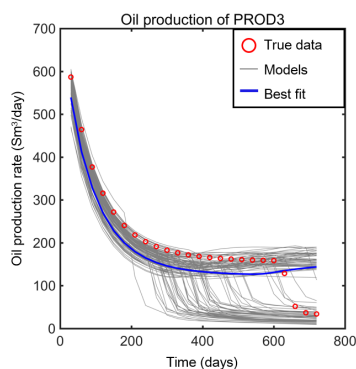
양상블칼만필터의 최종 모델은 생산정에서의 물돌파(water breakthrough) 시점이 참조모델과 크게 다른 것을 볼 수 있다. 이는 양상블 칼만필터의 기본가정인 정규분포를 따르지 않는 물성분포의 저류층에서는 제대로 교정이 이루어지지 못하는 양상블 칼만필터의 한계점을 보여준다. NSGA-II의 최종 모델도 참조모델에 비하여 물돌파 시점을 제대로 예측하지 못하였다. 이는 목적함수가 많은 경우에, 모든 목적함수를 만족시키는 모델을 찾기가 어려운 NSGA-II의 한계점을 보여준다. 반면에, 제안한 기법의 최종모델은 매칭을 수행한 기간인 720일까지 모든 생산정에서의 물돌파 시점을 잘 예측하였다. 720일 이전에 물돌파가 나타난 PROD3에서는 모든 모델이 720일 전에 물돌파가 나타났으며, 720일 이전에는 물돌파가 나타나지 않았다가 720일 이후에 물돌파가 나타난 PROD2의 물돌파 시점을 대부분의 모델이 예측하였다.



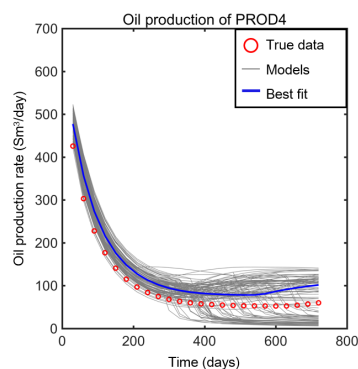
(a)



(b)

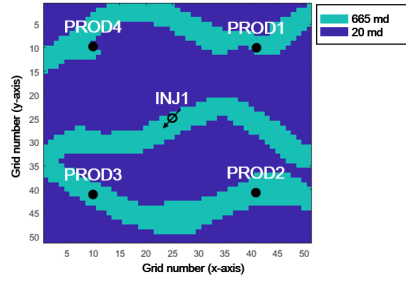


(c)

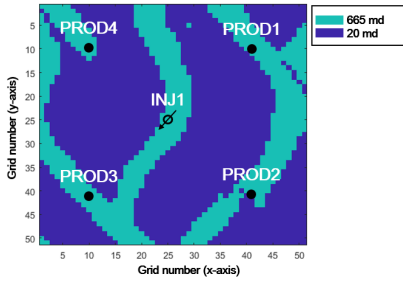


(d)

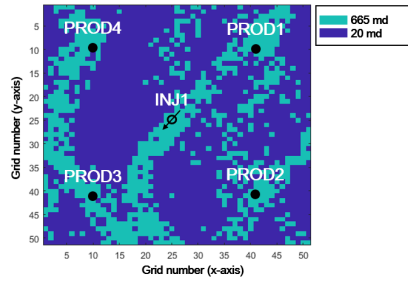
Figure 4.22 Oil production profiles of initial models until 720 days
(case 2): (a) PROD1, (b) PROD2, (c) PROD3, (d) PROD4.



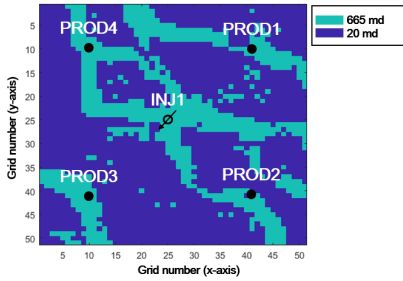
(a)



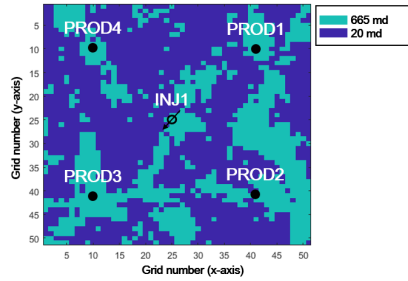
(b)



(c)



(d)



(e)

Figure 4.23 Permeability distribution of best fit model (case 2):

(a) reference field, (b) initial models,

(c) NSGA-II, (d) Ensemble Kalman-filter, (e) proposed method.

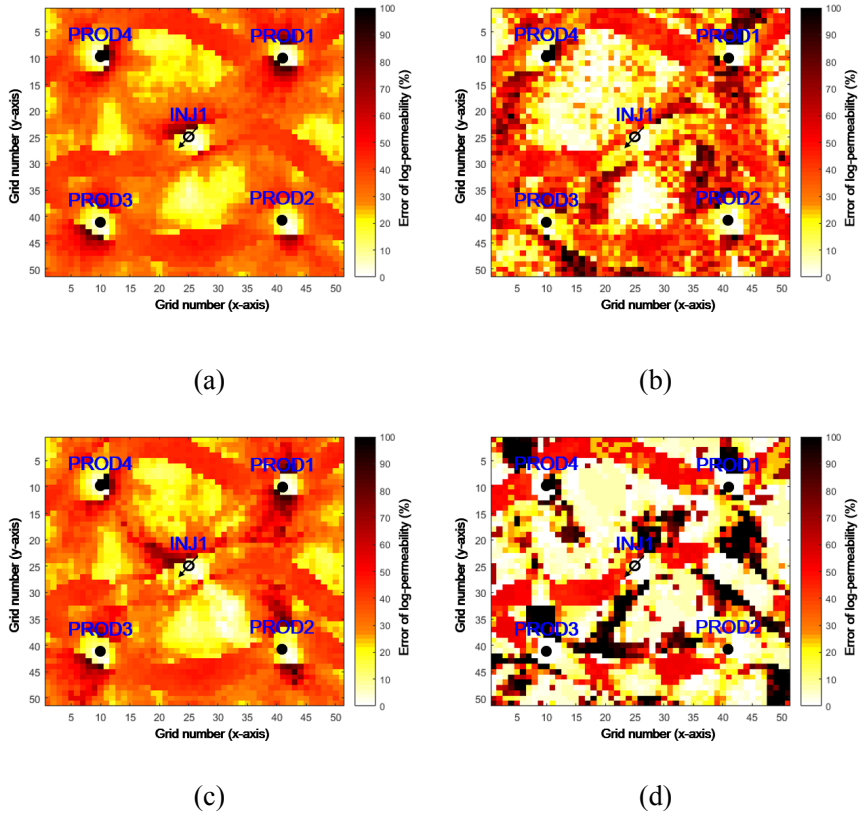
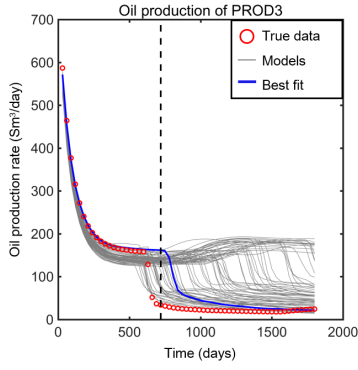
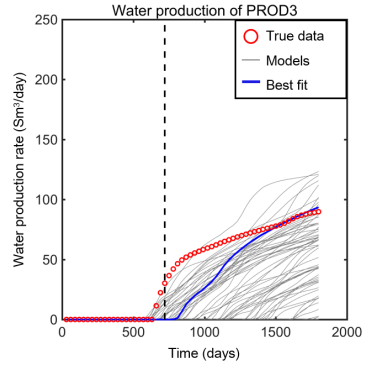


Figure 4.24 Average error map of permeability (case 2):

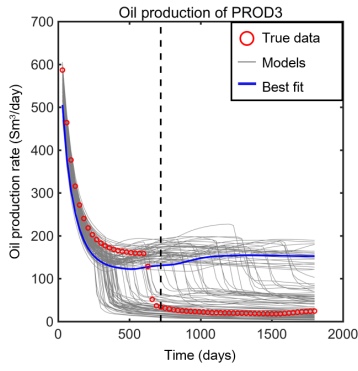
- (a) initial models, (b) matched models by NSGA-II, (c) assimilated models by Ensemble Kalman-filter, (d) selected models by proposed method.



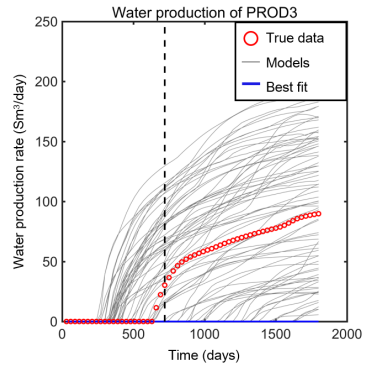
(a)



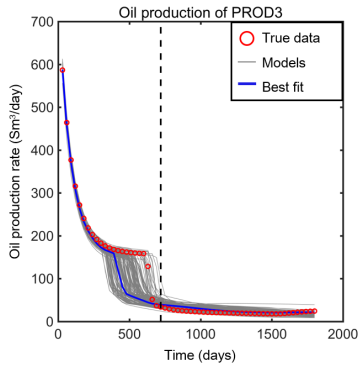
(b)



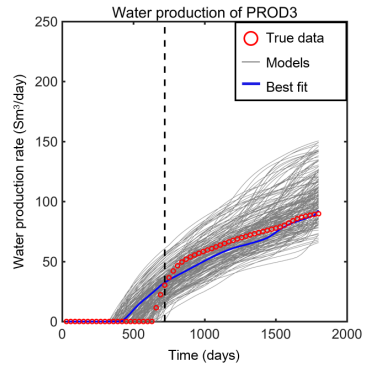
(c)



(d)



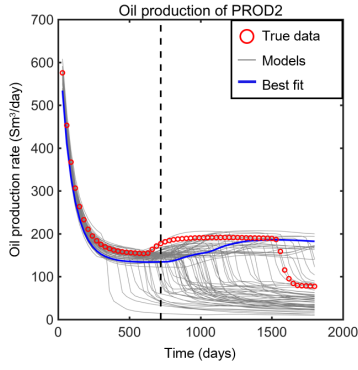
(e)



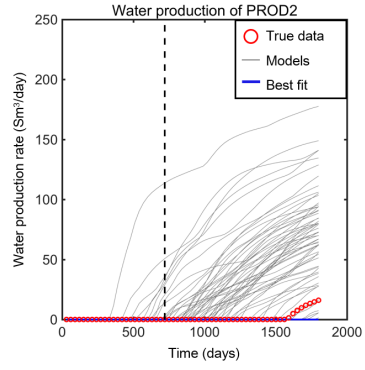
(f)

Figure 4.25 Production profiles of PROD3 (case 2):

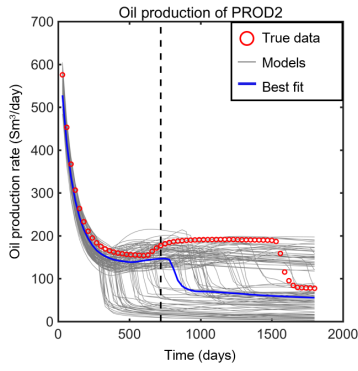
(a) (b) NSGA-II (oil, water production), (c) (d) Ensemble Kalman-filter (oil, water production), (e) (f) proposed method (oil, water production).



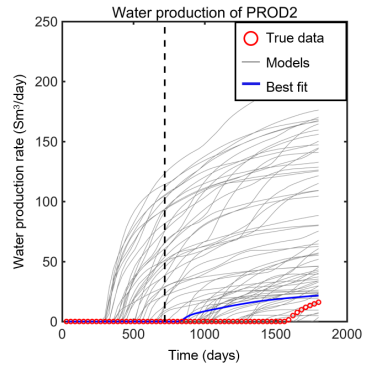
(a)



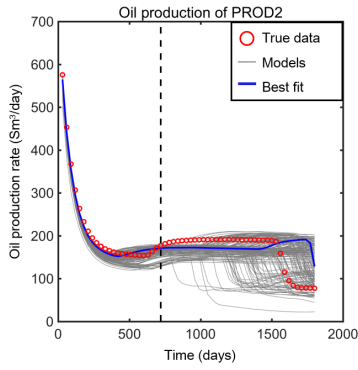
(b)



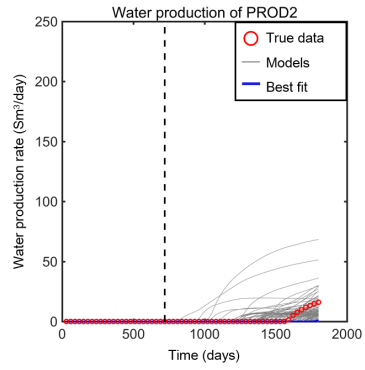
(c)



(d)



(e)



(f)

Figure 4.26 Production profiles of PROD2 (case 2):

(a) (b) NSGA-II (oil, water production), (c) (d) Ensemble Kalman-filter (oil, water production), (e) (f) proposed method (oil, water production).

Table 4.14는 각 기법을 통해 히스토리매칭을 수행한 결과들의 매칭구간 생산정별 평균 오일 생산량 오차를 요약한 표이다. 제안한 기법은 매칭구간에서의 각 생산정의 물돌과 시점을 잘 예측하였기 전체적으로 작은 오차를 보여주었으며, 앙상블칼만필터는 초기모델에 비하여 오차가 크게 개선되지 못한 모습을 보여주었다. Table 4.15는 각 기법을 통해 히스토리매칭을 수행한 결과들의 예측구간 생산정별 평균 오일 생산량 오차와 총 오일 생산량, 총 물 생산량 오차를 요약한 표이다. 앙상블칼만필터는 매칭구간과 마찬가지로 예측구간에서도 초기모델에 비하여 크게 개선되지 못하는 모습을 보여주었으나, NSGA-II는 초기모델이 비하여 오차가 어느 정도 감소하는 모습을 보여주었다. 하지만, PROD3의 물돌과 시점을 제대로 예측하지 못하여 PROD3에서의 생산량 오차가 매우 크게 나타났다. 제안한 기법은 모든 생산정의 물돌과 시점을 비교적 정확하게 예측함으로써 매칭구간과 마찬가지로 예측구간에서도 가장 낮은 오차를 보였다. 한편, 예측구간의 생산량 오차가 매칭구간에 비하여 크게 증가한 것을 알 수 있는데, 이는 초기모델의 트레이닝 이미지가 참조모델과 전혀 다르기 때문에 발생한 문제로 분석된다.

Figure 4.27은 각 기법을 통해 히스토리매칭을 수행한 결과들의 총 오일생산량과 총 물생산량을 나타낸 그래프이며, Figure 4.28은 박스플롯으로 나타낸 그림이다. 초기모델과 앙상블 칼만필터는 총 오일 생산량을 과소예측하고, 총 물 생산량을 과대예측하는 경향을 보여주면서 예측에 실패한 모습을 보여주었다. NSGA-II와 제안한 기법은 참조모델의 총 오일 생산량, 총 오일생산량을 예측범위 안에 포함하는 모습을 보여주었으며, 제안한 기법이 NSGA-II 보다 더 참조모델과 유사한 생산량을 예측하였다. 이상의 결과를 통하여 이 연구에서 제안한 기법은 저류층의 물성분포가 정규분포를 따르지

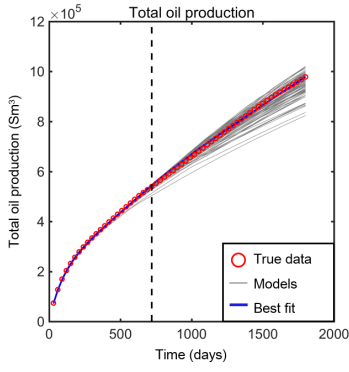
않는 채널저류층에서 제대로 된 교정이 이루어지지 못한 양상블
칼만필터와 달리 암상분포를 신뢰성 있게 추정하는 히스토리매칭
결과를 산출함을 확인하였다.

Table 4.14 Errors of history matching results for case 2
(during matching period, 720 days)

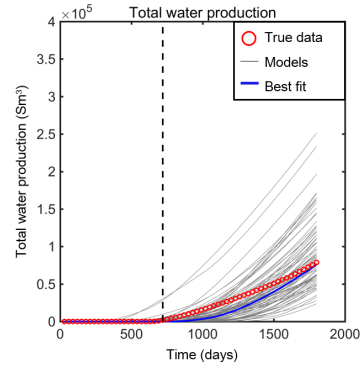
method	Initial models	NSGA-II	Ensemble Kalman-filter	Proposed method
PROD1	28.2%	12.2%	21.6%	4.2%
PROD2	27.2%	7.6%	20.9%	9.9%
PROD3	43.1%	35.4%	41.3%	10.1%
PROD4	50.3%	18.3%	50.3%	28.5%

Table 4.15 Errors of history matching results for case 2
(during prediction period, 1,800 days)

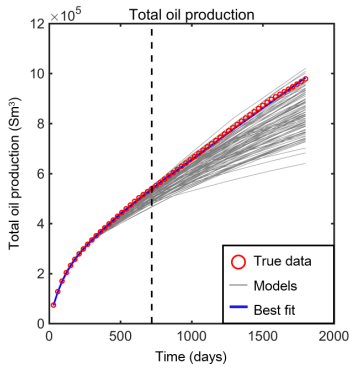
method	Initial models	NSGA-II	Ensemble Kalman-filter	Proposed method
PROD1	46.9%	33.6%	40.1%	26.9%
PROD2	47.3%	32.0%	39.6%	23.7%
PROD3	183.6%	202.4%	178.6%	14.9%
PROD4	59.1%	36.8%	59.5%	42.2%
Total oil production	7.0%	1.4%	5.1%	0.8%
Total water production	1568.9%	185.7%	1066.4%	76.4%



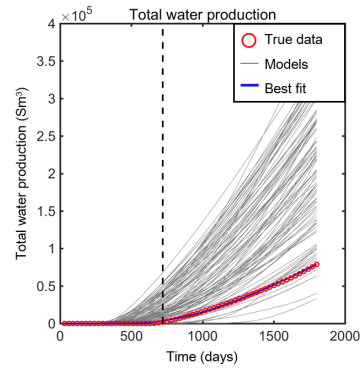
(a)



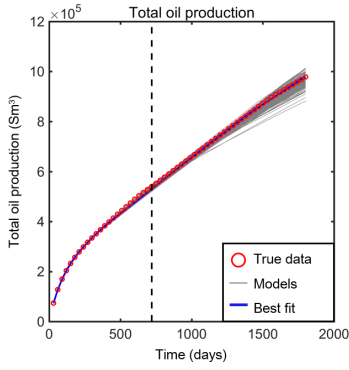
(b)



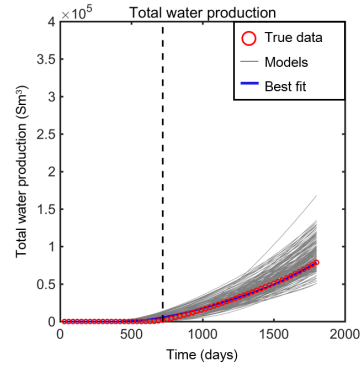
(c)



(d)



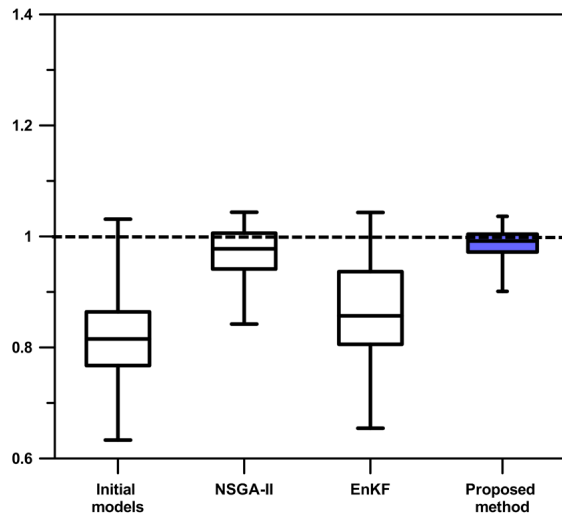
(e)



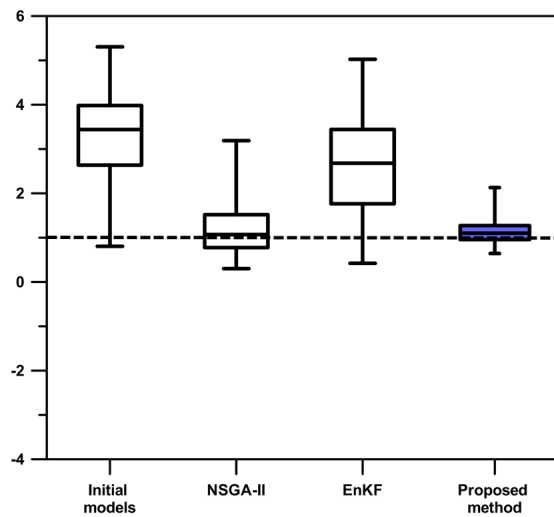
(f)

Figure 4.27 Total production (case 2):

- (a) (b) NSGA-II (total oil, water production),
(c) (d) Ensemble Kalman-filter (total oil, water production),
(e) (f) proposed method (total oil, water production).



(a)



(b)

Figure 4.28 Box plot of history matching results at 1,800 days (case 2):
 (a) Total oil production, (b) Total water production.

4.2 현장 적용성 평가

현장 적용성 평가를 위하여 이 연구에서 히스토리매칭을 수행한 PUNQ(Production forecasting with Uncertainty Quantification)-S3 필드는 Elf Exploration Production社가 운영하였던 실제 필드를 기반으로 만들어진 가상 필드이다. 유럽의 여러 회사와 대학들이 협업하여 개발하였으며 다양한 저류층 특성화 및 생산 최적화 연구에서 표준 테스트모델로서 널리 사용되고 있다(Egberts et al., 2002).

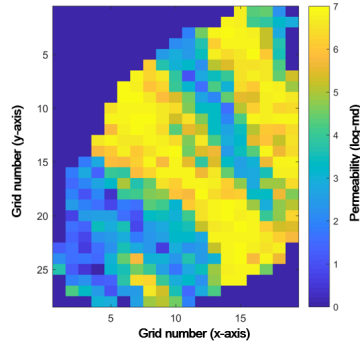
PUNQ-S3 필드는 $19 \times 28 \times 5 = 2,660$ 개의 격자를 가지고 있다. 수직방향으로 총 5개의 층으로 이루어졌으며, Table 4.16에서 볼 수 있듯이 층마다 유체투과도와 공극률의 분포범위가 다르다. 1, 3, 5층이 서로 비슷한 특성을 가지며 유체투과도와 공극률의 범위가 높아 생산성이 좋다. 2층은 유체투과도와 공극률의 범위가 낮아 생산성이 좋지 못하며, 4층은 2층과 1, 3, 5층의 중간 정도의 범위를 가진다. Figure 4.29는 PUNQ-S3의 각 층별 로그-유체투과도 분포도와 생산정의 위치를 나타낸다. 1, 3, 5층은 유사한 채널모양의 높은 유체투과도 분포를 보이며, 2층과 4층은 상대적으로 유체투과도가 낮다. 1층과 2층에는 생산정의 천공이 이루어지지 않았으며, 3층에는 PROD3, PROD4, 4층에는 6개 전부, 5층에는 PROD3, PROD2, PROD4의 천공이 이루어졌다.

히스토리매칭에 사용한 초기모델 100개는 PUNQ-S3의 참조필드와 동일하게 정규분포를 따르도록 지구통계학적 기법으로 생성되었다. Figure 4.30-31는 참조필드의 생산이력과 초기모델 100개의 생산량을 도시한 그래프이다. 빨간색 동그라미가 참조필드의 생산이력을 의미하며, 회색 선은 초기모델들의 생산량을 의미한다. 오일 생산량, 정저압력, 가스-오일비, 워터컷(watercut) 모두

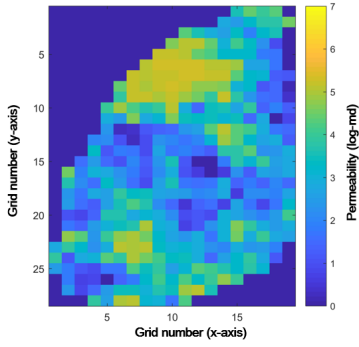
초기모델들의 생산량 범위가 넓게 나타나며 참조필드의 생산이력과의 오차가 크다는 것을 알 수 있다.

Table 4.16 Permeability and porosity range for the layers in PUNQ-S3 field.

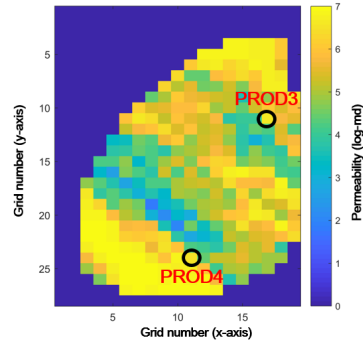
Layers	1, 3, 5	2	4	Unit
Porosity	0.01-0.30	0.01-0.17	0.01-0.22	-
Horizontal Permeability	0.1-1000	0.1-200	0.1-500	mD
Vertical Permeability	0.1-500	0.1-50	0.1-100	mD



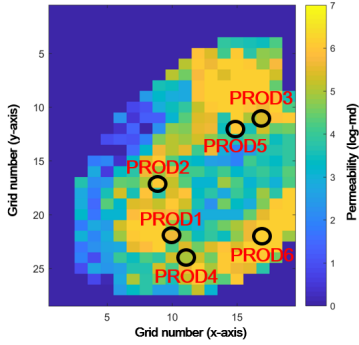
(a)



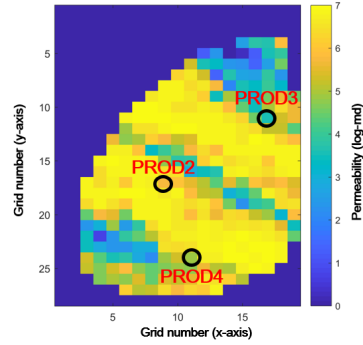
(b)



(c)

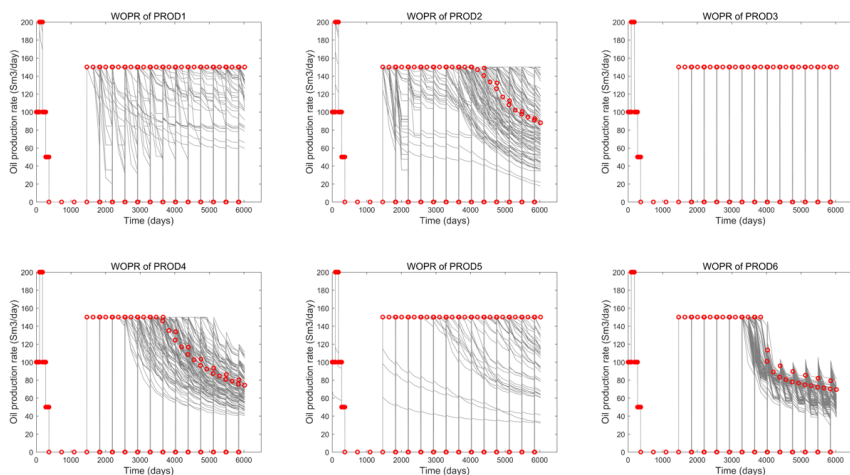


(d)

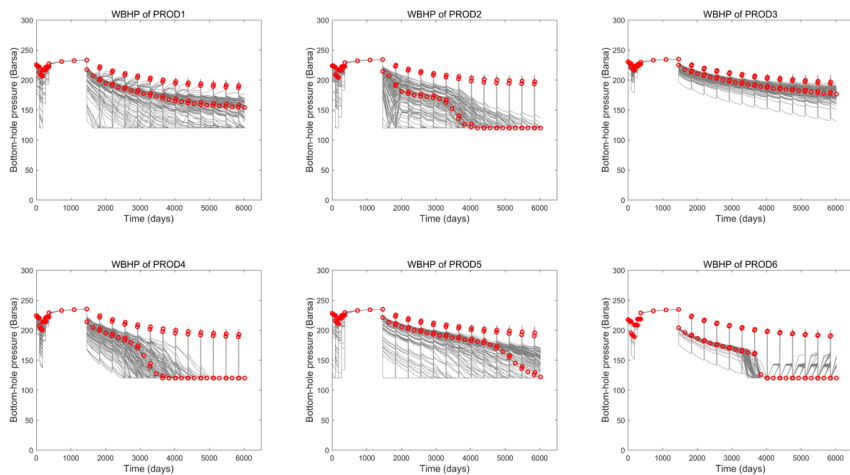


(e)

Figure 4.29 Log-permeability distribution of PUNQ-S3:
(a) layer 1, (b) layer 2, (c) layer 3, (d) layer 4, (e) layer 5.



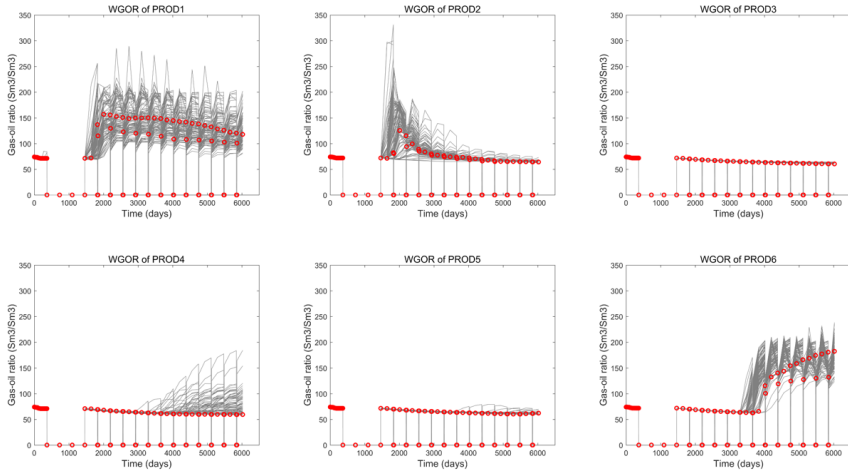
(a)



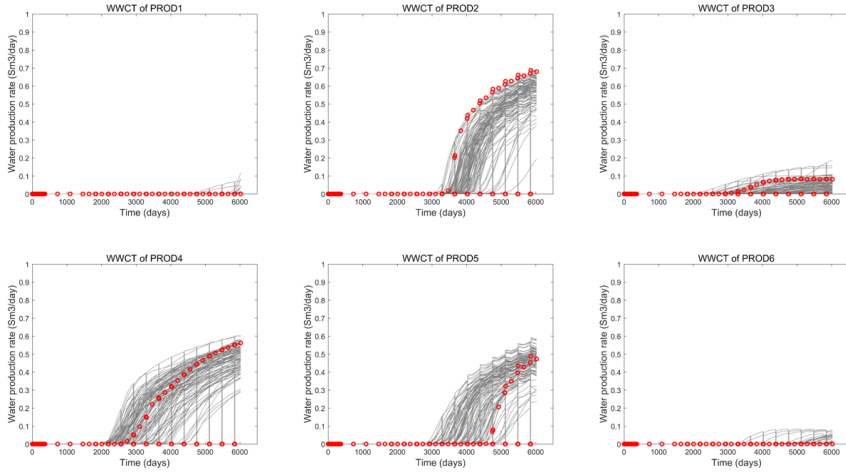
(b)

Figure 4.30 Comparison between production profiles of initial models and reference for PUNQ-S3 field:

(a) oil production rate, (b) bottom-hole pressure.



(a)



(b)

Figure 4.31 Comparison between production profiles of initial models and reference for PUNQ-S3 field:

(a) gas-oil ratio, (b) watercut.

4.2.1 심층신경망 은닉층 선정

제안한 기법으로 히스토리매칭을 수행하기에 앞서 $19 \times 28 \times 5 = 2,660$ 개 격자의 유체투과도를 100개의 출력층 뉴런으로 압축하는 심층 오토인코더를 구축하였다. 심층신경망의 안정성을 확보하기 위하여 두 번의 부호화 과정을 거쳤으며, 첫 번째 부호화 은닉층의 뉴런수에 대한 민감도 분석을 통하여 최적의 뉴런수를 선정하였다.

Table 4.17은 첫번째 부호화 은닉층 뉴런수를 각각 200개, 300개, 500개, 1,000개로 설정하였을 때의 학습오차와 연산시간을 비교한 표이다. 뉴런의 수가 많을수록 첫 번째 학습오차는 줄어들지만, 연산시간과 두 번째 학습오차가 증가하는 경향을 보였다. Figure 4.32는 첫 번째 부호화 은닉층의 뉴런수에 따른 총 학습시간과 총 학습오차를 나타낸 그래프이다. 첫 번째 부호화 은닉층의 뉴런수가 200개일 때, 가장 학습오차가 작고, 학습에 소요되는 시간도 짧은 것을 알 수 있다. 이 연구에서는 첫 번째 부호화 은닉층의 최적 뉴런수를 200개로 선정하였고, Table 4.18은 PUNQ-S3 필드의 히스토리매칭을 위하여 구축한 심층신경망의 입출력 변수를 나타낸다. Table 4.19는 심층신경망 기법을 통해 부호화된 출력층을 이용하여 구축한 PUNQ-S3 필드의 신경망 입출력 변수를 나타낸다.

Table 4.17 Training error and training time depending on the number of hidden neurons for PUNQ-S3 field

Number of output neurons for first encoding	200	300	500	1,000
Training error for first encoding	1.3%	0.9%	0.8%	0.9%
Training time for first encoding	358.6 sec	537.1 sec	972.8 sec	1823.6 sec
Training error for second encoding	1.6%	2.4%	4.0%	3.5%
Training time for second encoding	169.3 sec	225.6 sec	308.1 sec	620.3 sec

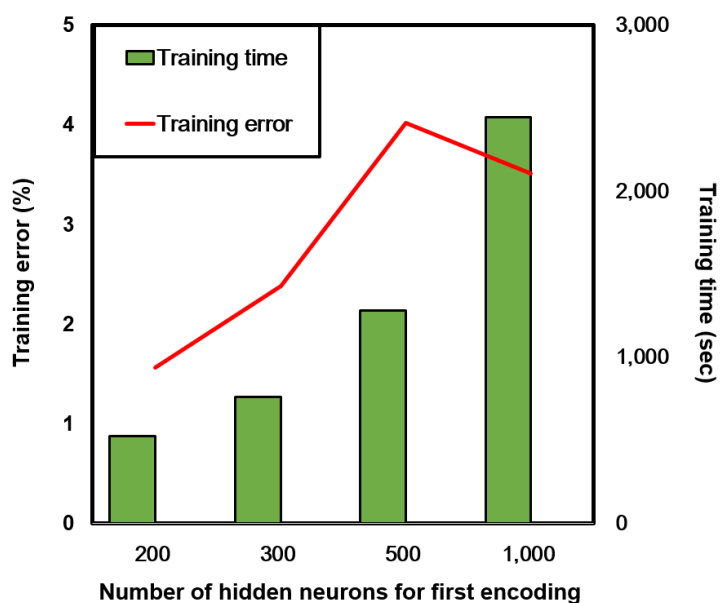


Figure 4.32 Training error and training time depending on the number of hidden neurons for PUNQ-S3 field.

Table 4.18 Neuron description of autoencoder for PUNQ-S3 field

Description of the neuron	Number of neurons and algorithm
Input neuron (Grid permeability)	$19 \times 28 \times 5 = 2660$
Hidden neuron	[200, 100]
Train algorithm	Trainscg†

† Scaled conjugate gradient backpropagation

Table 4.19 Neuron description of neural network for PUNQ-S3 field

Description of the neuron	Number of neurons and algorithm
Input neuron (oil rate, water rate)	$20 \times 6 \times 3 = 360$
Hidden neuron [Multi-layer]	[500, 500]
Output neuron (Encoded grid permeability)	100
Train algorithm	Traingdx††

†† Gradient descent with momentum and adaptive learning rate backpropagation

4.2.2 히스토리매칭 결과 비교분석

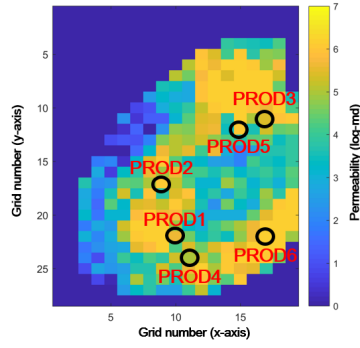
제안한 기법의 성능을 평가하기 위하여 제안한 기법을 통해 히스토리매칭을 수행한 결과와 NSGA-II, 앙상블칼만필터의 결과를 비교하였다. 제안한 기법과 동일한 초기모형을 사용하였으며, 동적자료도 제안한 기법과 동일한 360개의 동적자료를 사용하여 교정을 수행하였다. Schlumberger社의 ECLIPSE 100 소프트웨어를 통하여 저류층 시뮬레이션을 수행하였다.

Figure 4.33은 참조모델과 각 히스토리매칭 기법으로 생성된 모델들 중 총 오일 생산량 오차가 가장 작은 모델의 유체투과도 분포도를 보여준다. 모든 생산정의 천공이 이루어진 층이 4층이기 때문에 4층의 유체투과도 분포도를 비교하였다. 이 유체투과도 모델을 각 기법을 통해 히스토리매칭한 대표모델로 볼 수 있다. Figure 4.34는 각 기법으로 매칭된 최종 모델들과 참조모델의 평균 유체투과도 오차를 나타낸다. 하얀색에 가까울수록 오차가 낮아서 매칭이 잘된 부분이며, 까만색에 가까울수록 오차가 커서 매칭이 잘 안된 부분이다. 제안한 기법이 다른 기법에 비하여 밝은 부분이 훨씬 많이 보이며 상대적으로 유체투과도 예측이 잘 이루어졌음을 확인할 수 있다.

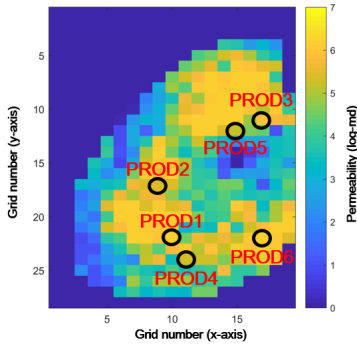
Figure 4.35-37는 초기모델과 각 기법을 통해 수행한 히스토리매칭 결과를 비교한 그림이다. 예측구간에서 물돌파가 나타난 생산정인 PROD2의 오일 생산량, 정저압력, 가스-오일비 그래프를 보여준다. 빨간색 동그라미는 참조모델의 생산이력을 의미하며, 회색선은 각 히스토리매칭 기법을 통해 만들어진 모델들의 생산량을 의미한다. 파란색 선은 Figure 4.33에 나타난 각 기법의 대표모델의 생산량이다. 세로로 된 검정색 점선은 마지막 히스토리매칭 구간인 2,936일을

나타낸다.

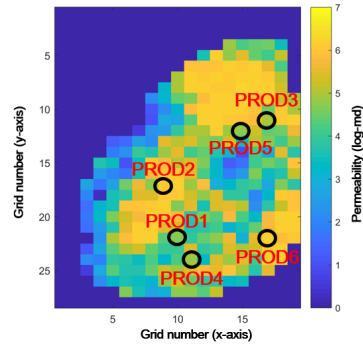
NSGA-II의 최종 모델은 초기모델에 비하여 크게 개선되지 못하는 예측결과를 보여주었다. 이는 PUNQ-S3 필드처럼 목적함수가 18개로 많은 경우에 모든 목적함수를 만족시키는 모델로 수립하기가 어려운 NSGA-II의 한계점을 보여준다. 앙상블칼만필터는 초기모델과 비교하여 참조필드의 생산이력과 유사한 생산량 결과를 산출하였다. 오일생산량과 정저압력, 가스-오일비 모두에서 초기모델에 비하여 작은 오차로 참조모델의 생산이력을 예측하였다. 이를 통해 앙상블칼만필터가 PUNQ-S3 필드와 같이 저류층 물성이 정규분포를 따르는 저류층에서는 신뢰성 있는 예측결과를 보여줄을 확인하였다. 제안한 기법은 다른 두 기법들에 비하여 작은 오차로 참조 모델의 생산이력을 예측하였으며, 특히 PROD2의 경우 참조모델과 오차가 거의 없는 생산이력을 보여주었다.



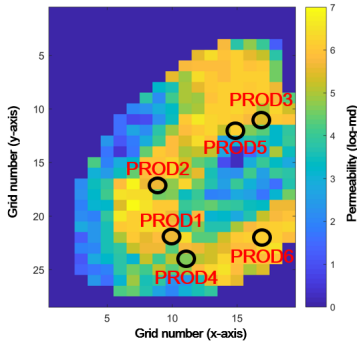
(a)



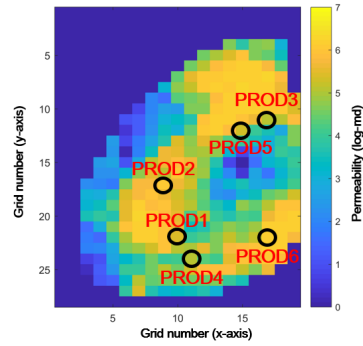
(b)



(c)



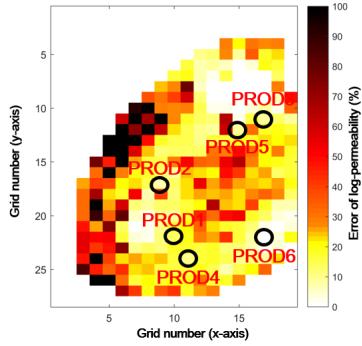
(d)



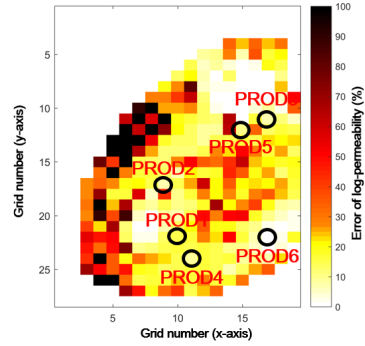
(e)

Figure 4.33 Log permeability distribution of best fit model for layer 4:

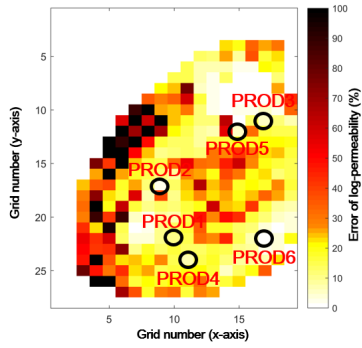
(a) reference field, (b) initial models, (c) NSGA-II,
(d) Ensemble Kalman-filter, (e) proposed method.



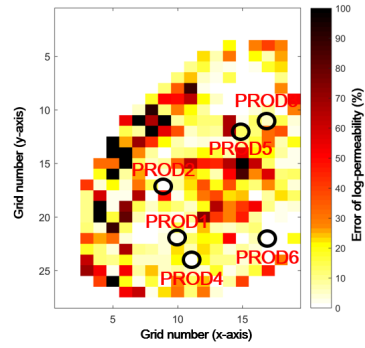
(a)



(b)

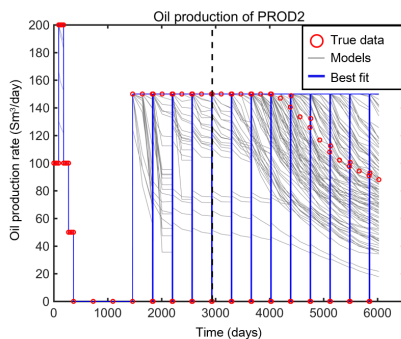


(c)

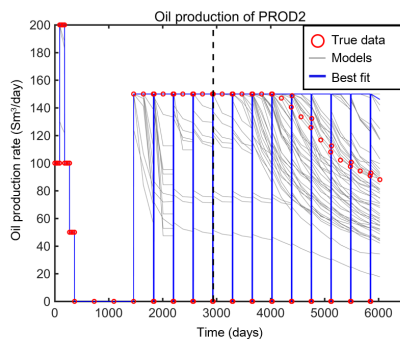


(d)

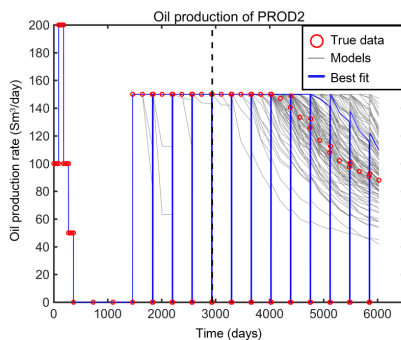
Figure 4.34 Average error map of log-permeability for layer 4:
(a) initial models, (b) NSGA-II, (c) Ensemble Kalman-filter,
(d) proposed method.



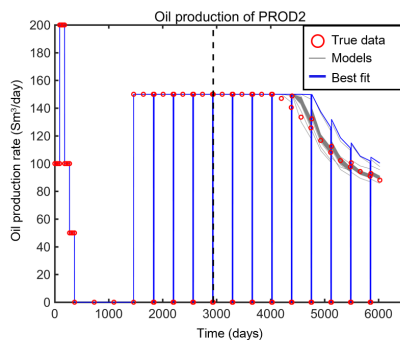
(a)



(b)



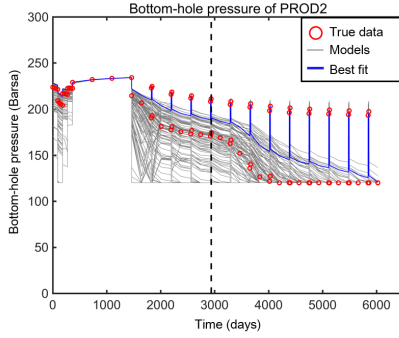
(c)



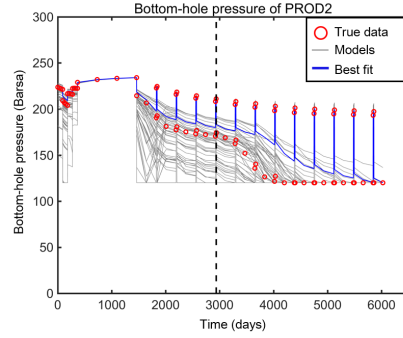
(d)

Figure 4.35 Oil production rate of PROD2:

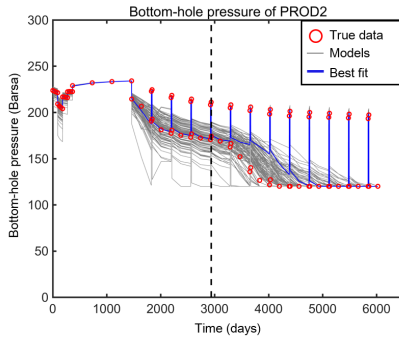
(a) initial models, (b) NSGA-II, (c) Ensemble Kalman-filter,
(d) proposed method.



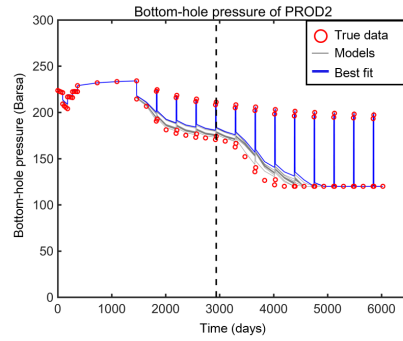
(a)



(b)



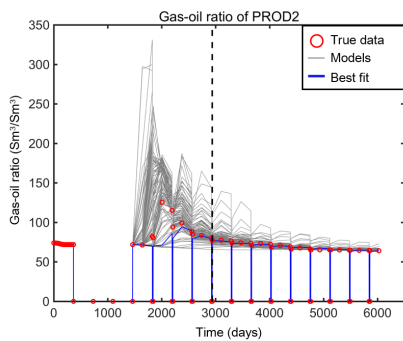
(c)



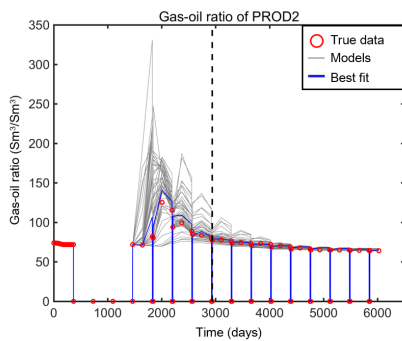
(d)

Figure 4.36 Bottom-hole pressure of PROD2:

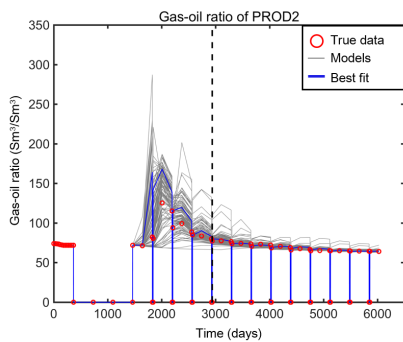
(a) initial models, (b) NSGA-II, (c) Ensemble Kalman-filter,
(d) proposed method.



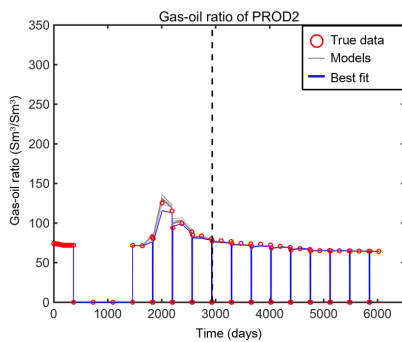
(a)



(b)



(c)



(d)

Figure 4.37 Gas-oil ratio of PROD2:

(a) initial models, (b) NSGA-II, (c) Ensemble Kalman-filter,
(d) proposed method.

Table 4.20-4.23은 각 기법을 통해 히스토리매칭한 결과들의 예측구간 생산정별 오일 생산량, 정저압력, 가스-오일비 평균오차를 요약한 표이다. NSGA-II는 목적함수가 너무 많아서 초기모델에 비하여 오차가 개선되지 않았으며, 대부분의 생산정에서 제안한 기법이 가장 작은 오차를 보임을 확인할 수 있다. Figure 4.38, 39는 각 기법을 통해 히스토리매칭을 수행한 결과들의 총 오일, 가스 생산량을 나타낸 그래프이며, Figure 4.40은 각 기법을 통해 히스토리매칭을 수행한 결과들의 마지막 예측일 총 오일, 가스 생산량을 박스플롯으로 나타낸 그림이다.

총 오일 생산량의 경우 초기 모델이 전체적으로 과소 예측하면서 오차가 컸는데, NSGA-II는 초기모델과 유사한 범위를 보여주며 예측이 실패했음을 알 수 있다. 앙상블칼만필터와 제안한 기법은 초기모델에 비하여 참조모델의 총 오일생산량에 가까운 결과를 산출하였다. 총 가스 생산량의 경우 NSGA-II는 초기모델의 범위와 크게 다르지 않는 범위를 보여준 반면에, 앙상블칼만필터와 제안한 기법은 2분위수가 참조모델의 총 가스생산량과 거의 일치하게 생산량을 예측하였다. 이상의 결과를 통하여 이 연구에서 제안한 기법은 목적함수가 18개인 불균질 필드에서 제대로 된 교정이 이루어지지 못한 NSGA-II와 달리 신뢰성 있는 히스토리매칭 결과를 산출함을 확인하였다.

Table 4.20 Errors of history matching results (oil production rate)

method	Initial models	NSGA-II	Ensemble Kalman-filter	Proposed method
PROD1	2.6%	1.9%	0.8%	0.0%
PROD2	6.8%	6.8%	3.4%	0.5%
PROD3	0.0%	0.0%	0.0%	0.0%
PROD4	5.3%	4.6%	3.4%	1.5%
PROD5	2.5%	2.7%	0.4%	0.5%
PROD6	3.9%	3.5%	3.3%	4.5%

Table 4.21 Errors of history matching results (bottom-hole pressure)

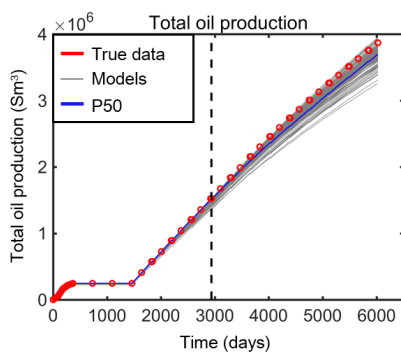
method	Initial models	NSGA-II	Ensemble Kalman-filter	Proposed method
PROD1	5.3%	4.2%	2.9%	0.4%
PROD2	4.8%	4.7%	2.7%	1.8%
PROD3	2.0%	1.4%	0.8%	1.4%
PROD4	3.3%	3.1%	2.3%	0.8%
PROD5	5.8%	5.7%	3.3%	5.0%
PROD6	2.2%	2.1%	1.6%	2.4%

Table 4.22 Errors of history matching results (gas-oil ratio)

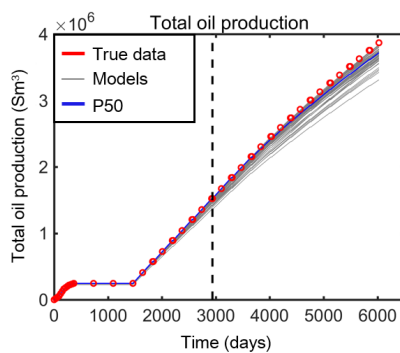
method	Initial models	NSGA-II	Ensemble Kalman-filter	Proposed method
PROD1	7.7%	8.4%	6.2%	1.1%
PROD2	4.4%	4.2%	3.2%	0.6%
PROD3	0.6%	0.6%	0.3%	0.3%
PROD4	4.2%	4.1%	2.6%	2.7%
PROD5	0.7%	0.7%	0.5%	0.1%
PROD6	5.9%	5.4%	5.5%	6.7%

Table 4.23 Errors of history matching results
(total oil and gas production)

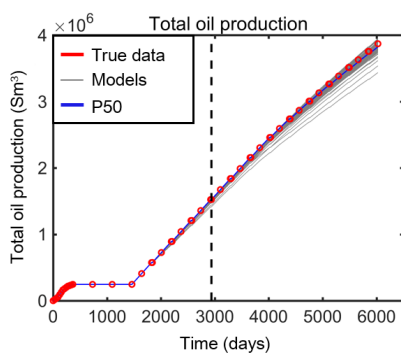
method	Initial models	NSGA-II	Ensemble Kalman-filter	Proposed method
Total oil production	4.7%	4.6%	1.3%	1.3%
Total gas production	2.3%	1.5%	0.6%	0.2%



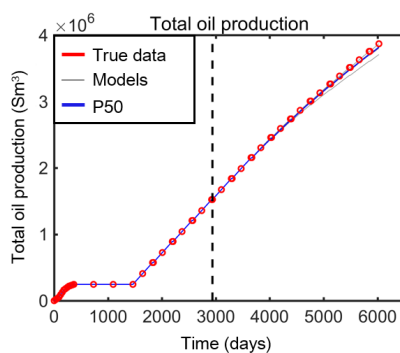
(a)



(b)

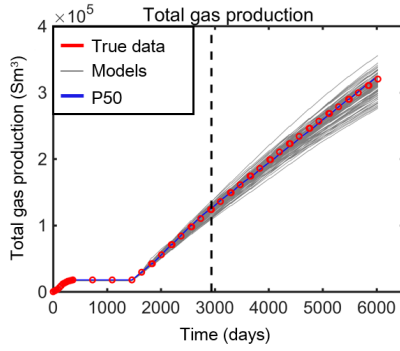


(c)

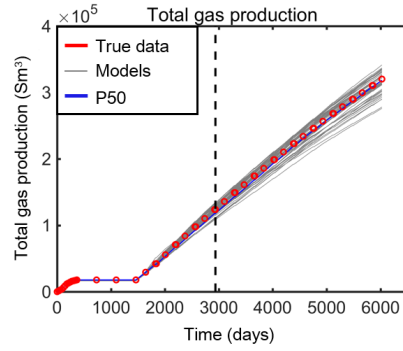


(d)

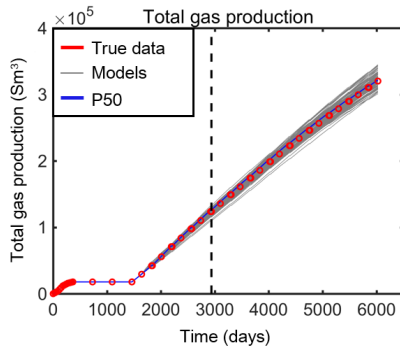
Figure 4.38 Total oil production of history matching results:
 (a) initial models, (b) NSGA-II, (c) Ensemble Kalman-filter,
 (d) proposed method.



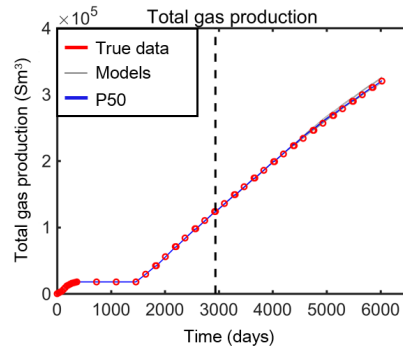
(a)



(b)

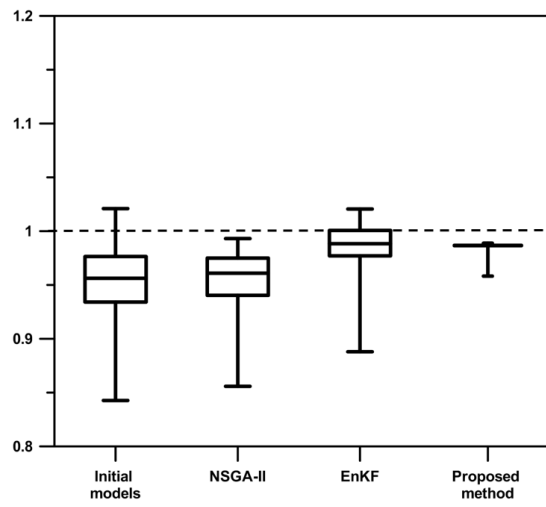


(b)

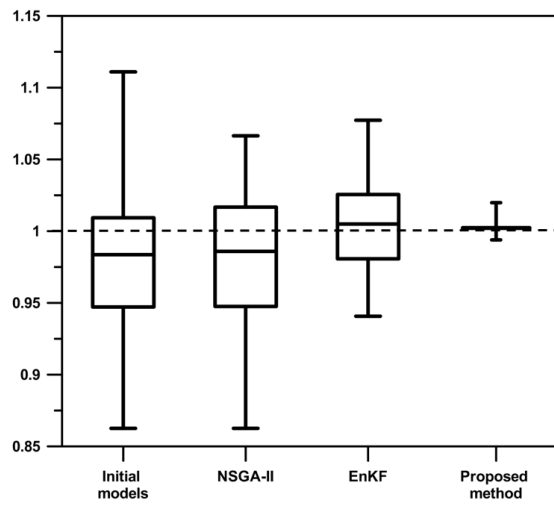


(d)

Figure 4.39 Total gas production of history matching results:
 (a) initial models, (b) NSGA-II, (c) Ensemble Kalman-filter,
 (d) proposed method.



(a)



(b)

Figure 4.40 Box plot of history matching results at 6,025 days:
(a) total oil production, (b) total gas production.

5 결론

이 연구에서는 심층신경망과 거리기반 후보군 선정기법을 결합하여 저류층의 생산 히스토리매칭을 수행하고 신뢰도 높게 생산추이를 예측하는 기법을 개발하였다. 저류층의 생산추이를 입력층, 저류층의 물성을 출력층으로 한 역산모델 신경망을 구성하고, 거리기반 후보군 선정을 통해 실제 생산자료와 유사한 추이를 보이는 후보 저류층을 새로운 학습집단으로 활용함으로써 신경망 기법의 학습성능을 개선하였다. 심층신경망을 이용하여 신경망의 출력층을 부호화함으로써 학습효율을 높였다.

이 연구에서 제안한 기법을 세일저류층, 채널저류층, 불균질 필드에 적용하여 다음과 같은 결론을 도출하였다.

(1) 심층신경망 기반의 다목적 히스토리매칭 기법을 개발하였다.

기존 연구들에서 신경망 기법을 프록시모델로 활용한 것과 달리 저류층의 생산량을 입력층, 저류층 물성을 출력층으로 한 역산모델 신경망을 구축함으로써 목적함수 의존적 문제를 극복하였다.

(2) 목적함수가 가스 생산량 하나이며, 지질학적 복잡성이 거의 없는 균질한 세일저류층에서 2년간 가스 생산이력에 대한 히스토리매칭을 통하여 제안한 기법의 방법론을 검증하였다. 거리기반 군집화를 바탕으로 한 거리기반 후보군 선정 기법을 통하여 새로운 학습집단을 선정함으로써 세대가

지날수록 신경망 기법의 학습성능이 개선됨을 확인하였다.

(3) 불균질 저류층들의 히스토리매칭을 수행하기에 앞서, 심층신경망을 이용하여 신경망의 출력층을 부호화함으로써 학습효율을 높였다. 각각의 저류층에서 민감도 분석을 수행하여 학습시간과 학습오차를 고려한 최적의 심층신경망 뉴런수를 선정하였다. 채널저류층에서는 출력층 뉴런수가 300개일 때, 학습시간대비 학습오차가 작았다. 3차원 필드인 PUNQ-S3 필드에서는 출력층 뉴런수가 200개일 때, 학습시간과 학습오차가 가장 작았다. 이와 같이 저류층 조건에 따라서 다른 출력층 뉴런수를 선정함으로써 최적의 신경망학습 결과를 산출하였다.

(4) 저류층의 암상분포가 생산량에 큰 영향을 주며, 정규분포를 따르지 않는 채널저류층에서 암상분포를 추정하였다. 4개의 생산정의 2년간 오일 생산량, 물 생산량 등 총 8개의 목적함수에 대한 다목적 히스토리매칭을 수행하였다. 역산한 최적화 저류층을 이용하여 향후 3년간의 생산이력을 예측하였고, 제안한 기법을 NSGA-II, 앙상블칼만필터와 비교하였다. 앙상블칼만필터는 기본가정인 정규분포를 만족시키지 않는 채널저류층에서 히스토리매칭에 실패한 반면에, 제안한 기법은 암상분포를 신뢰성 있게 추정함으로써 채널저류층의 생산량을 작은 오차로 예측하였다.

(5) 제안한 기법의 필드 적용성을 평가하기 위하여 6개의

생산정이 존재하는 PUNQ-S3 필드에서 오일 생산량, 정저압력, 가스-오일비 등 총 18개의 목적함수에 대한 다목적 히스토리매칭을 수행하였으며, 제안한 기법과 NSGA-II, 앙상블칼만필터를 비교하였다. 목적함수가 많아지면 최적해가 수렴하지 못하는 문제가 발생하는 NSGA-II는 히스토리매칭에 실패하였다. 반면에, 제안한 기법은 역산모델 신경망을 통하여 목적함수 의존적 문제를 극복하여 유체투과도 분포도 참조모델과 유사하게 모사하면서 생산량도 크게 줄였다.

이상의 연구결과를 통하여 제안한 거리기반 후보군 선정과 심층신경망을 결합한 다목적 히스토리매칭 기법이 목적함수 의존적 문제를 극복하고, 저류층의 물성분포와 상관없이 신뢰성 있는 최적해를 도출함을 확인하였다. 제안한 기법은 다수의 생산정이 존재하는 현장 저류층에서 정확한 생산량 예측이 가능하여 향후 저류층 운영계획 수립을 위한 핵심적 도구로 활용할 수 있다.

참고문헌

안성인, 2017. *심층 오토인코더를 결합한 신경망 기법의 채널저류층 특성화 연구*, 서울대학교 공학석사학위논문.

이경복, 2014. *거리기반 앙상블스무더를 이용한 채널저류층 특성화*, 서울대학교 공학박사학위논문.

임서진, 2016. *거리기반방법과 앙상블칼만필터를 이용한 암종모델 히스토리매칭 연구*, 서울대학교 공학석사학위논문.

Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C. and Vallès, B., 2009. "The ensemble Kalman filter in reservoir engineering-a review," *SPE Journal*, 14(3): 393–412.

Anderson, D.M., Liang, P. and Okouma, V., 2012. "Probabilistic forecasting of unconventional resources using Rate Transient Analysis: case studies," presented at the SPE Americas Unconventional Resources Conference, Pittsburgh, PA, USA, 5–7 June.

Araghinejad, S., 2013. *Data-Driven Modeling: Using MATLAB in Water Resources and Environmental Engineering*, Springer Science & Business Media.

- Ballester, P.J. and Carter, J.N., 2007. "A parallel real-coded genetic algorithm for history matching and its application to a real petroleum reservoir," *Journal of Petroleum Science and Engineering*, 59(3): 157–168.
- Bengio, Y., Lamblin, P., Popovici, D., and Larochelle, H., 2007. *Greedy layer-wise training of deep networks*, Advances in neural information processing systems.
- Berkhin, P., 2006. "A survey of clustering data mining techniques," *Grouping multidimensional data*, 25: 71.
- Centilmen, A., Ertekin, T. and Grader, A. S., 1999. "Applications of neural networks in multiwell field development," presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, USA, 3–6 October.
- Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T.A.M.T., 2002. "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE transactions on evolutionary computation*, 6(2): 182–197.
- Doraisamy, H., Ertekin, T. and Grader, A.S., 1998. "Key parameters controlling the performance of neuro-simulation applications in field development," presented at the SPE Eastern Regional Meeting, Pittsburgh, Pennsylvania, USA, 9–11 November.

Egberts, P.J.P., Brouwer, G.K. and Bos, C.F.M., 2002. "History matching and forecasting with uncertainty quantification; a real case study," presented at the EAGE 64th Conference & Exhibition, Florence, Italy, 27–30 May.

Evensen, G., 1994. "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research: Oceans*, 99(C5): 10143–10162.

Evensen, G., Hove, J., Meisingset, H.C., Reiso, E., Seim, K.S. and Espelid, Ø., 2007. "Using the enKf for assisted history matching of a north sea reservoir model," presented at the SPE Reservoir Simulation Symposium, Houston, Texas, 26–28 February.

Gu, Y. and Oliver, D.S., 2005. "History matching of the PUNQ-S3 reservoir model using the ensemble Kalman filter," *SPE journal*, 10(02): 217–224.

Gu, Y. and Oliver, D.S., 2006. "The ensemble Kalman filter for continuous updating of reservoir simulation models, *Journal of Energy Resources Technology*, 128(1): 79–87.

Han, Y., Park, C. and Kang, J.M., 2011. "Prediction of nonlinear production performance in waterflooding project using a multi-objective evolutionary algorithm," *Energy Exploration & Exploitation*, 29(2):

129–142.

Hartigan, J.A. and Wong, M.A., 1979. “Algorithm AS 136: A k-means clustering algorithm,” *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1): 100–108.

Hinton, G., Osindero, S. and Teh, Y., 2006. “A Fast Learning Algorithm for Deep Belief Nets,” *Neural Computation*, 18(7): 1527–1554.

Holland, J.H., 1975. *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*, MIT press.

Kalman, R.E., 1960. “A new approach to linear filtering and prediction problems,” *Journal of basic Engineering*, 82(1): 35–45.

Kaufman, L. and Rousseeuw, P.J., 1987. “Clustering by means of medoids,” In *Statistical data analysis based on the L1-norm and related methods*, North-Holland, 405–416.

Kim, J., Kang, J.M., Park, C., Park, Y., Park, J. and Lim, S., 2017. “Multi-objective history matching with a proxy model for the characterization of production performances at the shale gas reservoir,” *Energies*, 10(4): 579.

- Kim, J., Kang, J.M., Park, Y., Lim, S., Park, C., and Park, J., 2016. “Probabilistic estimation of shale gas reserves implementing Fast Marching Method and Monte Carlo simulation” presented at the ASME 2016 35th International Conference on Ocean, Offshore and Arctic Engineering, Busan, Korea, 191–24 June.
- Kim, K., Ju, S., Ahn, J., Shin, H., Shin, C. and Choe, J., 2015. “Determination of key parameters and hydraulic fracture design for shale gas productions,” presented at the Twenty-fifth International Ocean and Polar Engineering Conference, ISOPE, Kona, HI, USA, 21–26 June.
- Lee, K., Jeong, H., Jung, S. and Choe, J., 2013. “Improvement of ensemble smoother with clustered covariance for channelized reservoirs,” *Energy Exploration & Exploitation*, 31(5): 713–726.
- Li, D. and Yu, D., 2014. “Deep learning: methods and applications,” *Foundations and Trends® in Signal Processing*, 7 (3–4): 197–387.
- Maltarollo, V.G., Honório, K.M. and da Silva, A.B.F., 2013. “Applications of artificial neural networks in chemical problems,” In *Artificial neural networks-architectures and applications*, InTech.
- Min, B., Park, C., Kang, J.M., Park, H.J. and Jang, I., 2011. “Optimal Well-placement Based on Artificial Neural Network Incorporating the Productivity Potential,” *Energy Sources, Part A*, 33(18): 1726–11738.

- Morales-German, G., Navarro-Rosales, R. and Dubost, F.X., 2012. “Production forecasting for shale gas exploration prospects based on statistical analysis and reservoir simulation,” presented at SPE Latin America and Caribbean Petroleum Engineering Conference, Mexico City, Mexico, 16–18 April.
- Nævdal, G. and Vefring, E.H., 2002. “Near-well reservoir monitoring through ensemble Kalman filter,” presented at the SPE/DOE Improved Oil Recovery Symposium, Tulsa, Oklahoma, USA, 13–17 April.
- Oliver, D.S. and Chen, Y., 2011. “Recent progress on reservoir history matching: a review,” *Computational Geosciences*, 15(1): 185–221.
- Park, H., Lim, J. and Kang, J.M., 2005. “Analysis of multi-well gas reservoir performance using Polynomial Neural Network (PNN) with layer overpassing structure,” *Journal of the Korean Society for Geosystem Engineering*, 42(3): 152–159.
- Qian, N., 1999. “On the momentum term in gradient descent learning algorithms,” *Neural networks*, 12(1): 145–151.
- Schmidhuber, J., 2015. “Deep learning in neural networks: An overview,” *Neural networks*, 61: 85–117.
- Seber, G.A., 2009. *Multivariate observations*, John Wiley & Sons.

Shahkarami, A., Mohaghegh, S.D., Gholami, V. and Haghighat, S.A., 2014.

“Artificial Intelligence (AI) assisted history matching,” presented at the SPE Western North American and Rocky Mountain Joint Meeting, SPE, Denver, Colorado, USA, 17–18 April.

Velmurugan, T. and Santhanam, T., 2010. “Computational complexity between k-means and k-medoids clustering algorithms for normal and uniform distributions of data points,” *Journal of Computer Science*, 6(3): 363–368.

Xie, J., Yang, C., Gupta, N., King, M.J. and Datta-Gupta, A., 2015. “Depth of investigation and depletion in unconventional reservoirs with Fast-Marching Methods,” *SPE Journal*, 20(4): 831–841.

Appendices

A. 인공신경망과 심층 오토인코더의 성능비교

인공신경망과 심층 오토인코더의 학습성능을 비교하기 위하여, 채널저류층과 3차원 필드를 대상으로 인공신경망과 심층 오토인코더를 한번만 수행한 결과를 비교하였다. 매칭구간에서의 각 생산정의 생산량 오차와 학습에 소요된 시간을 비교하였다.

Table A.1-A.2는 채널저류층에서 매칭구간 동안의 생산량 오차와 학습에 소요된 시간을 비교한 표이다. 두 가지 신경망기법 모두 초기모델에 비하여 개선된 오차를 보여줬으며, 유사한 예측성능을 보여주었다. 하지만, 학습시간에서는 심층 오토인코더가 기존 인공신경망기법에 비하여 약 8배 정도 빠른 결과를 보여주었다.

Table A.3-A.6은 PUNQ-S3 필드에서 매칭구간 동안의 생산량 오차를 비교한 표이다. 두 가지 신경망기법 모두 초기모델에 비하여 개선된 오차를 보여줬으며, 총 오일생산량, 총 가스생산량, 18개의 목적함수 오차의 평균 모두 심층 오토인코더가 더 작은 오차를 보여주었다. 학습시간에서도 심층 오토인코더가 기존 인공신경망기법에 비하여 약 3.5배 이상 빠른 결과를 보여주었다.

Table A.1 Comparison of history matching results between neural network and proposed method for channelized reservoir (oil production rate)

method	Initial models	Neural network	Proposed method
PROD1	28.2%	7.6%	13.9%
PROD2	27.2%	12.4%	11.3%
PROD3	43.1%	37.6%	39.1%
PROD4	50.3%	32.2%	27.3%

Table A.2 Comparison of history matching results between neural network and proposed method for channelized reservoir (total)

method	Initial models	Neural network	Proposed method
Training time	-	2424.0 sec	319.52 sec
Mean of error	37.2%	22.5%	22.9%

Table A.3 Comparison of history matching results between neural network and proposed method for PUNQ-S3 field (oil production rate)

method	Initial models	Neural network	Proposed method
PROD1	1.4%	0.5%	0.1%
PROD2	1.8%	0.5%	0.1%
PROD3	0.0%	0.0%	0.0%
PROD4	0.1%	0.0%	0.0%
PROD5	0.7%	0.0%	0.1%
PROD6	0.0%	0.0%	0.0%

Table A.4 Comparison of history matching results between neural network and proposed method for PUNQ-S3 field (bottom-hole pressure)

method	Initial models	Neural network	Proposed method
PROD1	3.8%	3.2%	1.4%
PROD2	5.3%	3.4%	2.1%
PROD3	1.4%	0.5%	1.5%
PROD4	3.7%	2.0%	1.4%
PROD5	3.8%	1.4%	4.9%
PROD6	0.5%	2.7%	0.2%

Table A.5 Comparison of history matching results between neural network and proposed method for PUNQ-S3 field (gas-oil ratio)

method	Initial models	Neural network	Proposed method
PROD1	5.2%	3.2%	4.1%
PROD2	6.9%	5.6%	3.5%
PROD3	0.2%	0.1%	0.1%
PROD4	0.4%	0.3%	0.2%
PROD5	0.4%	0.1%	0.3%
PROD6	0.2%	0.2%	0.1%

Table A.6 Comparison of history matching results between neural network and proposed method for PUNQ-S3 field (total)

method	Initial models	Neural network	Proposed method
Total oil production	2.0%	0.6%	0.1%
Total gas production	1.3%	2.7%	1.1%
Training time	-	2196.9 sec	625.6 sec
Mean of error	2.0%	1.3%	1.1%

B. 유전 알고리즘

유전 알고리즘(Genetic Algorithm)은 Holland(1975)가 제안한 최적화 기법 중 하나로 자연세계의 진화과정을 이용한 계산 모델이다. 생물의 진화를 모방하는 대표적인 진화기반 기법으로, 선택(selection), 변이(mutation), 교차(crossover), 인구(population), 세대(generation) 등 실제 진화과정에서 많은 부분을 차용하였다. 유전 알고리즘은 최적화 문제에서 초기 세대의 해들에 대해서 선택, 교차, 변이의 과정을 거쳐 점점 더 좋은 해를 찾아간다. Figure B.1은 유전 알고리즘의 순서도를 나타낸다. 유전 알고리즘은 다음과 같은 순서를 통하여 최적해를 찾아간다.

- (1) 주어진 최적화 문제에 적합한 범위 내에서 모두 포함할 수 있는 초기 세대를 임의로 구성한다.
- (2) 다음 세대로 전해지는 해의 후보가 되는 해들을 선택한다.
일반적인 유전 알고리즘에서는 목적함수 값을 최소화하는 해들을 선택한다.
- (3) 선택된 해들의 유전자(구성인자)를 조합하여 새로운 해들을 만들어 낸다. 일반적으로 임의의 두 해를 선택하여 교차 연산을 수행하게 되며, 부모 해에서 겹치지 않는 유전인자를 받아 새로운 유전자를 구성한다.

- (4) 변이를 통하여 추가로 새로운 해들을 만들어 낸다. 변이 연산은 임의의 해의 유전 인자의 순서 혹은 값이 임의로 변경되어 다른 해로 변형되는 연산이다. 변이를 통하여 지역 최적해에 빠지는 경우를 방지할 수 있으며, 해의 다양성을 높여준다.
- (5) 설정한 정지조건(세대수, 목적함수 값 등)을 만족할 때까지 (2)~(4) 과정을 반복한다.

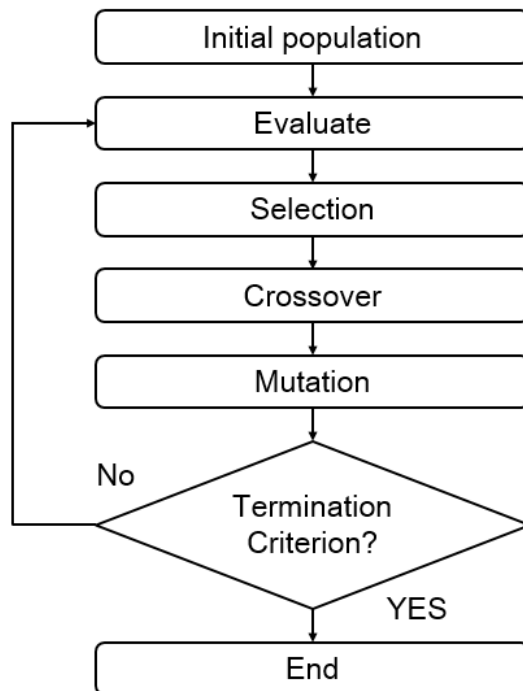


Figure B.1 Flow chart of Genetic Algorithm.

C. 앙상블칼만필터

Kalman(1960)은 관측오차를 이용하여 모델의 인자들을 교정하는 칼만필터(Kalman filter)를 개발하였다. 칼만필터는 선형문제에만 적용이 가능하며 하나의 모델에 대해서만 교정과 예측을 수행하기 때문에, 불확실성 평가가 불가능하다는 단점이 있다. Evensen(1994)은 칼만필터를 개선하여 비선형문제에 적용할 수 있고, 등가의 확률을 가지는 다수의 모델인 앙상블(Ensemble)을 사용하여 불확실성 평가가 가능한 앙상블칼만필터를 제안하였다.

앙상블칼만필터는 앙상블들의 평균을 참으로 가정하고, 실제 관측값과의 추정오차공분산(estimated error covariance)을 최소화하는 방향으로 인자들을 교정한다. 앙상블칼만필터는 저류층의 정적인자, 동적인자, 시뮬레이션 결과로 구성된 상태벡터(state vector)를 통해 인자들을 교정함으로써 히스토리매칭을 수행한다. 정적인자는 격자별 유체투과도, 격자별 공극률 등과 같이 시간에 따라서 변하지 않는 물성을 의미하며, 동적인자는 격자별 압력, 물포화도와 같이 시간에 따라서 변하는 물성값을 의미한다. 시뮬레이션결과는 정적인자와 동적인자를 통해 수행한 시뮬레이션 결과로써 오일 생산량, 가스 생산량, 정저압력 등을 의미한다. Figure C.1은 앙상블칼만필터의 순서도를 나타낸다.

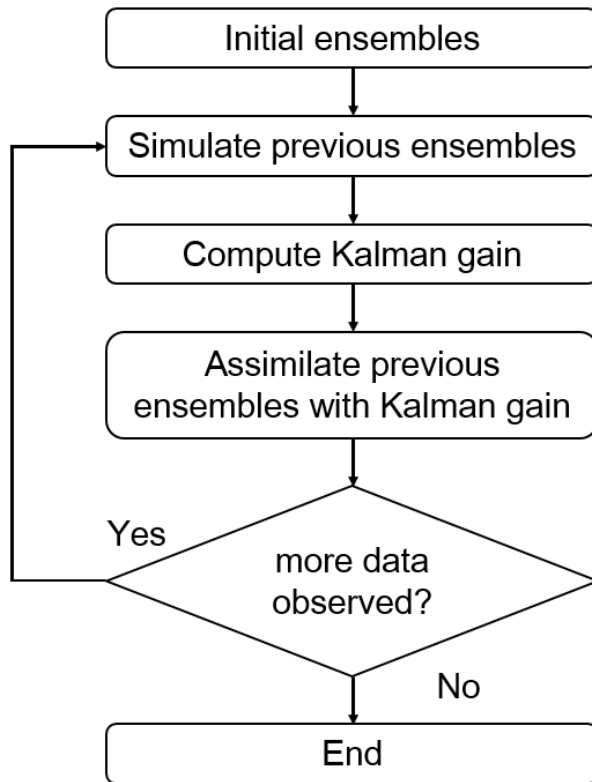


Figure C.1 Flow chart of Ensemble Kalman-filter.

총 앙상블의 개수를 N_e , 총 교정회수를 N_t 라고 하면, i 번째 앙상블에 대한 t 번째 상태벡터는 식 (C.2)와 같이 정의된다.

$$y_{t,i} = \begin{bmatrix} m_t^s \\ m_t^d \\ d_t \end{bmatrix}, i = 1, \dots, N_e, t = 1, \dots, N_t \quad (C.2)$$

여기에서 m_t^s 가 정적인자, m_t^d 가 동적인자, d_t 가 시물레이션 결과이다. 앙상블 칼만필터는 예측과 교정을 반복하면서 히스토리매칭을 수행한다.

예측단계에서는 d_t 를 계산하여 상태벡터를 구성한다. 첫 번째 예측단계에서는 정적자료를 이용하여 생성된 초기 저류층 모델인 $m_{1,i}^s$ ($i = 1, \dots, N_e$) 을 이용한다. 모든 앙상블들은 첫 관측자료가 얻어지는 시점까지만 저류층 시물레이션을 수행한다. 식 (C.3)와 같은 전위모델을 사용하여 예측을 하게 된다.

$$\begin{Bmatrix} m_t^d \\ d_t \end{Bmatrix} = f(m_{t-1}^s, m_{t-1}^d) \quad (C.3)$$

여기에서 f 는 다공질매질에서의 유체유동방정식을 의미하며 $t - 1$ 시점에서 t 시점까지의 저류층 시물레이션을 수행한다.

교정단계에서는 식 (C.4)와 같이 상태벡터를 교정한다. 칼만게인(Kalman gain) K 와 관측값 d_t^{obs} 와 예측값 Hy_t^p 의 차이를 곱한 값을 이전 시점의 상태벡터에 더하여 새로운 상태벡터를 계산한다. 새롭게 계산된다.

$$y_t^a = y_t^p + K(d_t^{obs} - Hy_t^p) \quad (C.4)$$

여기에서 y_t^a 는 교정 후의 상태벡터, y_t^p 는 교정 전의 상태벡터를 의미하며 H 는 측정행렬(measurement matrix operator)로써 상태벡터에서 예측값을 추출하는 행렬이다. 칼만게인 K 는 추정오차공분산을 최소로 하도록 식 (C.5)와 같이 계산된다.

$$K = C_Y^p H^T (H C_Y^p H^T + C_D)^{-1} \quad (C.5)$$

여기에서 T 는 전치행렬을 의미하며, C_Y^p 와 C_D 는 각각 교정전의 추정오차공분산과 관측오차공분산(measurement noise covariance)을 의미한다. 추정오차공분산은 공분산의 정의에 따라서 식 (C.6)과 같이 계산되며, 여기에서 추정오차 e_t 는 식 (C.7)로 정의된다.

$$C_{Y,t} = E[e_t e_t^T] = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} e_{t,i} e_{t,i}^T \quad (C.6)$$

$$e_t \equiv y_t - \hat{y}_t \quad (C.7)$$

여기에서 \hat{y}_t 는 참값을 의미한다. 앙상블칼만필터에서는 앙상블의 평균을 참값으로 가정하기 때문에, 초기 앙상블디자인이 히스토리매칭 결과에 큰 영향을 미친다(이경복, 2014).

Abstract

Multi-objective History Matching Integrated with Deep Neural Network and Distance-based Candidate Selection

Jaejun Kim

Department of Energy System Engineering

The Graduate School

Seoul National University

This study presents a multi-objective history matching method by integrating deep neural network and training reservoir models with a distance-based clustering. The conventional history matching methods tend to show divergence to a global optimality for non-Gaussian statistical assumptions and for the case of the many objective functions. To overcome these limitations, the proposed method includes the neural network as an inverse modelling tool and improves training performances through k-medoids clustering. Also, a deep autoencoder is included in the proposed neural network for the efficient data processing.

The proposed scheme is validated by showing the reliable predictability of gas production rate in a shale gas reservoir with bi-wing hydraulic fractures. The method is able to apply to the non-Gaussian permeability distribution of the channelized reservoir with a reasonable accuracy, whilst the ensemble Kalman-filter cannot be used. To examine applicability in the field, the proposed method has been compared with other methods for the heterogeneous reservoir with 18 objective functions. Though the optimal solution from multi-

objective evolutionary algorithm is not converged, the proposed method can estimate a reliable future productions by overcoming the objective dependent problem.

The proposed method can be a useful tool for decision-making in the field-scale reservoir development, since the approach developed in this study characterizes various reservoirs by estimating reservoir properties with reasonable accuracy.

Keywords: distance-based clustering, deep neural network, autoencoder, history matching, multi-objective optimization, reservoir characterization

Student Number: 2012-24169